Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/002633

International filing date: 28 January 2005 (28.01.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/607,684
Filing date: 07 September 2004 (07.09.2004)

Date of receipt at the International Bureau: 07 April 2005 (07.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

1300754

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

March 24, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/607,684*
FILING DATE: *September 07, 2004*
RELATED PCT APPLICATION NUMBER: *PCT/US05/02633*

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PTO/SB/16 (04-04)
Approved for use through 07/31/2006. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EV440868529US

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Mark Alan Gurkan	Ostermeier Guntas	Baltimore MD Baltimore MD

Additional inventors are being named on the _____ *separately numbered sheets attached hereto*

TITLE OF THE INVENTION (500 characters max)

Molecular Switches and Methods For Making and Using the Same Involving the Circular Permutation of DNA

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ Customer Number:

OR

☑ Firm or Individual Name	Johns Hopkins University				
Address	100 N. Charles Street				
Address	5th Floor				
City	Baltimore	State	MD	Zip	21201
Country	USA	Telephone	410-516-8300	Fax	410-516-5113

ENCLOSED APPLICATION PARTS *(check all that apply)*

☑ Specification *Number of Pages* 71 ☐ CD(s), Number _____

☐ Drawing(s) *Number of Sheets* _____ ☐ Other (specify) _____

☐ Application Data Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☑ Applicant claims small entity status. See 37 CFR 1.27.

☐ A check or money order is enclosed to cover the filing fees.

☐ The Director is hereby authorized to charge filing fees or credit any overpayment to Deposit Account Number: _____

☑ Payment by credit card. Form PTO-2038 is attached.

FILING FEE Amount ($): $80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☐ No.

☒ Yes, the name of the U.S. Government agency and the Government contract number are: GM066972

[Page 1 of 2]

Respectfully submitted,

SIGNATURE Rachel Cassidy

TYPED or PRINTED NAME Rachel Cassidy

TELEPHONE 410-516-8300

Date 07-Sept-04

REGISTRATION NO. 55,883 *(if appropriate)*

Docket Number: 4399

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

CERTIFICATE OF EXPRESS MAILING

EXPRESS MAILING LABEL NO.

EV440868529US

I hereby certify that this correspondence (along with any papers referred to as being attached or enclosed) is being deposited with the United States Postal Service as Express Mail, Post Office to Addressee with sufficient postage in a **Flat Rate** envelope addressed to MS Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450 on the date indicated below:

07-Sept-04
DATE of Signature
And of Mail Deposit

Signature

Johns Hopkins University RECEIVED

Licensing and Technology Development

BEST AVAILABLE COPY

Report of Invention Disclosure Form

This form is to be completed and submitted to the JHU office of Licensing and Technology Development (LTD) by anyone who believes they have developed a new invention. The purpose of this form is to enable LTD to evaluate whether legal protection to the invention will be sought and/or commercialization pursued. In order for this Report of Invention to be processed by LTD, it must be signed and dated by all inventors, and by the JHU Department Director(s) for all departments involved with the development of this invention. LTD can not process this report until it is complete with all necessary signatures found in Sections A, B and/or C. Visit the LTD web site at http://www.jhu.edu/technology/roi.html for HTML and Word downloadable formats of this form.

INVENTION INFORMATION

Title of Invention:
Molecular Switches and Methods for Making and Using the Same Involving the Circular Permutation of DNA

School(s) and Department(s) in which invention was developed: School of Engineering, Dept. of Chemical and Biomolecular Engineering

Additional inventors: ☒ Yes ☐ No If yes, please complete Additional Inventors section for each inventor.

Lead Inventor Information: ["The Lead Inventor is the primary contact person for LTD on all matters associated with this Report of Invention, including processing, patent prosecution and licensing. For reasons of administrative efficiency, it is the responsibility of the Lead Inventor to keep all other JHU inventors named on this Report of Invention informed of the status of such matters."]

Name of Lead Inventor: Ostermeier (Last) Marc (First) Alan (Middle) Ph.D. (Degree)

Title or Position:	School:	Department:
Assistant Professor	Engineering	Chemical and Biomolecular Engineering

Business phone:	Business fax:	E-mail:
410.516.7144	410.516.5510	oster@jhu.edu

Business address: Department of Chemical and Biomolecular Engineering, 221 Maryland Hall, 3400 N. Charles St., Baltimore, MD 21218

Interdepartmental address:

Home phone number: Home fax number:

Home address: 6916 Bellona Ave., Baltimore, MD 21212-1108

Citizenship: US Social Security Number:

Are you a Howard Hughes Medical Institute employee or investigator? ☐ Yes ☒ No

Are you a Kennedy Krieger Institute employee or investigator? ☐ Yes ☒ No

LTD Internal Use Only: REF 4399 TLA CGG Field of Use _____

ADDITIONAL INVENTOR(S)

Please copy this page for additional inventors as necessary

Name of Inventor: Guntas (Last) Gurkan (First) _____ (Middle) B.S. (Degree)

Title or Position: graduate student School: Engineering Department: Chemical and Biomolecular Engineering

Business phone: (410)516-4146 Business fax: (410)516-5510 E-mail: gurkanguntas@yahoo.com

Business address: Department of Chemical and Biomolecular Engineering, 221 Maryland Hall, 3400 N. Charles St., Baltimore, MD 21218
Interdepartmental address:

Home phone number: Home fax number:

Home address: 6230 Green Meadow Parkway, Baltimore, MD, 21209

Citizenship: Turkey Social Security Number:

Are you an Howard Hughes Medical Institute employee or investigator? ☐ Yes ☒ No
Are you a Kennedy Krieger Institute employee or investigator? ☐ Yes ☒ No

Name of Inventor: _____ (Last) (First) (Middle) (Degree)

Title or Position: School: Department:

Business phone: ( ) - Business fax: ( ) - E-mail:

Business address:
Interdepartmental address:

Home phone number: Home fax number:

Home address:

Citizenship: Social Security Number:

Are you an Howard Hughes Medical Institute employee or investigator? ☐ Yes ☐ No
Are you a Kennedy Krieger Institute employee or investigator? ☐ Yes ☐ No

INVENTION DESCRIPTION

Describe the invention completely, using the outline given below. Please provide an **Electronic Copy** of the invention disclosure document, references, and abstracts, in Windows format, on CD-Rom or Floppy Disk.

1. Abstract of the Invention [In order to assist Licensing and Technology Development with the assessment of this technology, please provide a summary of the invention that should be written to be understood by a wide audience including non-technical individuals]

The invention provides molecular switches which couple external signals to functionality and to methods of making and using the same. The switches according to the invention can be used, for example, to regulate gene transcription, target drug delivery to specific cells, transport drugs intracellularly, control drug release, provide conditionally active proteins, perform metabolic engineering, create molecular sensors, and modulate cell signaling pathways. Libraries comprising the switches and expression vectors and host cells for expressing the switches are also provided. These libraries involve the circular permutation of DNA.

2. Problem Solved [Describe the problem solved by this invention]

This invention enables the creation of fusions between nucleic acids by the insertion of one piece into another in order to functionally couple existing functionalities, modulate existing functionalities or create novel functionalities. The following list of applications that these engineered nucleic acids or proteins can have is illustrative of their broad potential: (a) regulation of gene transcription, (b) modulation of cell signaling pathways, (c) targeted drug delivery, (d) drug transport, (e) conditionally active toxic proteins, (f) metabolic engineering, (g) biosensors and (h) controlled drug release. Domain insertion is the most promising and most general strategy for engineering a molecular switch, as other strategies are limited in the sorts of signals that can be employed or the types of proteins that can be controlled. However, there are several limitations of existing domain insertion strategies that this invention overcomes: (1) they are not combinatorial. A combinatorial approach systematically tests a plurality of potential switches to find the optimum switch. (2) The switching behavior achieved by existing methods was generally modest (less than 2-fold effect). (3) None of the studies have explored circular permutation of either the target or the inserted gene; thus, existing domain insertion strategies can only explore a limited number of geometric configurations between the two domains.

3. Novelty [Identify those elements of the invention that are new when compared to the current state of the art]

To our knowledge, a circularly permuted gene that has been circularly permuted at a specific site and inserted into another gene (i.e. by domain insertion at a specific site or by random domain insertion) has never been described. To our knowledge, the insertion of a randomly circularly permuted gene into another gene (i.e. by domain insertion at a specific site or by random domain insertion) has never been described. Furthermore, the idea that such a process will create a molecular switch does not have precedent and is not obvious.

ACKNOWLEDGMENT, CERTIFICATION and ASSIGNMENT OF INVENTION

In order for this Report of Invention to be complete and processed by LTD, it must be signed and dated by:

(1) the JHU Department Director for each JHU department involved with the development of this invention (SECTION A), and,

(2) ALL Inventors (SECTIONS B and C), including those Inventors not subject to The Johns Hopkins University Intellectual Property Policy. Each Inventor must complete only one of Sections B or C (See explanations below).

(3) Please duplicate Sections A, B and/or C as needed for proper completion with ALL appropriate signatures.

SECTION A. JHU DEPARTMENT DIRECTOR'S ACKNOWLEDGEMENT

I have read and understood this Report of Invention.

JHU Department Director Signature	Michael Betenbaugh Typed or Printed Name Chem. + Biomol. Eng. JHU School / Department	Date
JHU Department Director Signature	Typed or Printed Name JHU School / Department	Date
JHU Department Director Signature	Typed or Printed Name JHU School / Department	Date

SECTION B. JHU INVENTOR CERTIFICATION and ASSIGNMENT

This section is to be completed only by those JHU personnel subject to The Johns Hopkins University Intellectual Property Policy. Non-JHU Inventors, and HHMI or KKI Inventors at JHU are subject to a separate assignment and must complete Section C. JHU Inventors who believe they are not subject to The Johns Hopkins University Intellectual Property Policy for the invention described herein must complete Section C.

I/we, the Inventors, hereby certify that the information set forth in this Report of Invention is true and complete to the best of my/our knowledge.

I/we, the Inventors, hereby certify that I/we will promptly advise LTD of any commercial interest regarding the invention described herein.

I/we, the Inventor(s), subject to The Johns Hopkins University Intellectual Property Policy and not under an obligation to assign intellectual property rights to another party, hereby affirm that in consideration for The Johns Hopkins University's evaluation of commercial potential and a share of income which I/we may receive upon commercialization of my/our invention, on the date of my/our signature(s) as indicated below do hereby assign and transfer my/our entire right, title and interest in and to the invention described herein unto The Johns Hopkins University, its successors, legal representatives and assigns.

JHU Inventor Signature	MARC OSTERMEIER Typed or Printed Name	Date
JHU Inventor Signature	GURKAN GUNTAS Typed or Printed Name	Date
JHU Inventor Signature	Typed or Printed Name	Date
JHU Inventor Signature	Typed or Printed Name	Date
JHU Inventor Signature	Typed or Printed Name	Date

4399

Molecular Switches and Methods for Making and Using the Same Involving the Circular Permutation of DNA

Abstract

The invention provides molecular switches which couple external signals to functionality and to methods of making and using the same. The switches according to the invention can be used, for example, to regulate gene transcription, target drug delivery to specific cells, transport drugs intracellularly, control drug release, provide conditionally active proteins, perform metabolic engineering, create molecular sensors, and modulate cell signaling pathways. Libraries comprising the switches and expression vectors and host cells for expressing the switches are also provided. These libraries involve the circular permutation of DNA.

Background

Molecular Switches. A hallmark of biological systems is the high degree of interactions amongst and within their constituent components. One advantage that such interactions bring is the establishment of coupling between different functions. A protein that couples two functions can be described as a molecular switch. For example, an allosteric enzyme is a switch that couples effector levels (input) to enzyme activity (output). In most general terms, a molecular switch couples signals (e.g. ligand binding, protein-protein interactions, pH, covalent modification, temperature) to functionality (e.g. enzymatic activity, binding affinity, fluorescence). Molecular switches can be of an "on/off" nature or such that the signal modulates the function between two different levels of activity. The network of such molecular switches establishes the complex circuits that control cellular processes.

The design of molecular switches to modulate or report on biological functions has enormous potential for a variety of applications including the creation of biosensors (Siegel and Isacoff 1997; Baird, Zacharias et al. 1999; Doi and Yanagawa 1999; de Lorimier, Smith et al. 2002; Fehr, Frommer et al. 2002) modulators of gene transcription and cell signaling pathways (Rivera 1998; Guo, Zhou et al. 2000; Picard 2000), and novel biomaterials (Stayton, Shimoboji et al. 1995) (see also Figure 6). Despite their vast potential, molecular switches have not been explored extensively, in part due to the paucity of universal strategies for engineering them and the difficulty in engineering a switch that responds to a signal unrelated to the proteins' function and activates in the presence of the signal (as opposed to, for example, active-site inhibitors). In general, existing strategies for creating switches are inherently limited in the nature of the function that can be controlled, the signal that can be employed, the lack or reversibility, the lack of sensitivity, or the requirement for additional cellular components.

Domain Insertion. Gene fusion technology, the fusion of two or more genes into a single gene, has been widely used as a tool in protein engineering, localization and purification. There are two conceptually different methods of making fusions. The simplest method of end-to-end fusions has been used almost exclusively. The second method is insertional fusion in which one gene is inserted into the middle of the other gene. Insertions result in a continuous domain being split into a discontinuous domain. In nature, although continuous domains are more common, discontinuous domains are not

rare (Russell and Ponting 1998) and a systematic survey of structural domains indicated that 28% of structural domains are discontinuous (Jones, Stewart et al. 1998).

The observation of insertions in naturally occurring proteins suggests that such a route can be viable to construct proteins with desired properties and function. Furthermore, insertional fusions offer an advantage in numbers, as there are many more insertional fusions that can be made between two proteins than end-to-end fusions. Whereas it is expected that many of these insertional fusions will not be able to fold correctly, studies on insertions of DNA coding for insertions of a few amino acids to entire proteins, summarized below, have succeeded in creating active, often fully active, hybrids.

Early work on one or two codon insertions indicated that such small insertions had a smaller effect on activity than larger insertions (Boeke 1981; Stone, Atkinson et al. 1984; Barany 1985; Barany 1985; Freimuth and Ginsberg 1986; Barany 1987). Work in Paul Schimmel's lab in the late 1980's showed that insertions of random sequences into *E. coli* methionyl-tRNA synthetase of up to 14 amino acids could be tolerated in select locations (Starzyk, Burbaum et al. 1989). Subsequently other large insertions (Zebala and Barany 1991; Ladant, Glaser et al. 1992; Hallet, Sherratt et al. 1997), even a randomized 120 amino acid library insertion (Doi, Itaya et al. 1997) have confirmed this plasticity of proteins for insertions.

The first example of successful insertion of one protein into another was of alkaline phosphatase (AP) into the E. coli outer membrane protein MalF, constructed as a tool for studying membrane topology (Ehrmann, Boyd et al. 1990). High levels of alkaline phosphatase activity were obtained in the fusions despite the fact that alkaline phosphatase requires dimerization for activity. Since then, AP has been successfully inserted into a number of integral membrane proteins (Bibi and Beja 1994; Lacatena, Cellini et al. 1994; Pigeon and Silver 1994; Sarsero and Pittard 1995; Pi and Pittard 1996; Cosgriff and Pittard 1997). Furthermore, other proteins with their N- and C-termini proximal, including green fluorescent protein GFP) (Siegel and Isacoff 1997; Biondi, Baehler et al. 1998; Kratz, Bottcher et al. 1999; Siegel and Isacoff 2000), TEM1 β-lactamase (Betton, Jacob et al. 1997; Doi and Yanagawa 1999; Collinet, Herve et al. 2000), thioredoxin (Lu, Murray et al. 1995), dihydrofolate reductase (Collinet, Herve et al. 2000), FKBP12 (Tucker and Fields 2001), estrogen receptor-α (Tucker and Fields 2001) and β-xylanase (Aÿ, Götz et al. 1998), have been successfully inserted into other proteins.

Relatively few studies have examined or attempted functional coupling in insertional fusions (i.e. creating a molecular switch). Although the goal of the work was not to functionally coupled the proteins, studies of the insertion of β-lactamase into MBP (Betton, Jacob et al. 1997) and the insertion of β-lactamase or DHFR into yeast phosphoglycerate kinase (PGK) (Collinet, Herve et al. 2000) indicate that a minimal level of functional coupling of the two activities can exist, even in fusions designed to minimize interaction between the two domains. More recently, yeast sensors for ligand binding were constructed by the insertion of FKBP12 and estrogen receptor-α ligand-binding domain (EPα-LBD) into a rationally chosen site in dihydrofolate reductase (DHFR) (Tucker and Fields 2001) in the first designed coupling of growth rate of an organism to a small molecule ligand by domain insertion. Yeast expressing the FKBP12-DHFR or ERα -DHFR fusion proteins had an approximate two-fold increase in growth

rate in the presence of their respective ligands (FK106 and estrogen) when DHFR activity limited growth. However, in vitro neither the switches activity nor stability significantly changed upon ligand binding and the mechanism for increased growth rate of cells bearing these genes is not clear.

Domain insertion for has been used to couple ligand binding and changes in fluorescence. The optical signal transduction of the green fluorescent protein (GFP) has made it an attractive target for engineering biosensors by domain insertion (Guerrero and Isacoff 2001). GFP has been inserted into voltage-gated channels for potassium (Siegel and Isacoff 1997) and sodium (Ataka and Pieribone 2002) to generate sensors in which voltage driven rearrangements in the channel alter the brightness of GFP by 5.0% and 0.5% respectively. Initial attempts at creating such molecular sensors by inserting β-lactamase into GFP (Doi and Yanagawa 1999) were unsuccessful; however, random mutagenesis on the fusion was able to create a protein whose fluorescence increased 60% upon binding of the β-lactamase inhibitory protein. Insertions of calmodulin (a Ca^{2+} binding protein) into GFP resulted in a fusion whose fluorescence changed up to 40% upon increases in Ca^{2+} concentration (Baird, Zacharias et al. 1999). In a related strategy, the gene for a circularly permuted GFP was sandwiched between the gene's for calmodulin and it's target peptide M13 to create a series of sensors whose fluorescence intensity increased, decreased or showed a excitation wavelength change upon binding $Ca2^{+}$ (Nagai, Sawano et al. 2001). The GFP sensors described in this paragraph were developed through trial and error. For example, the site for inserting GFP into the sodium channel that showed a 0.5% change was the only one of the eight sites tried that showed a response (Ataka and Pieribone 2002).

Domain insertion has also been used in a strategy called mutually exclusive folding (ref) in which either one or the other of the two domains can be folded at one time. A signal (such as the presence of the ligand of one the domains) can alter the equilibrium between which domain is folded (and active). The ligand can stabilize the domain it binds to, thus causing the other domain to unfold. However, such a system is an off-switch (the signal turns the function that is to be controlled off) and thus is not as generally useful for molecular switch applications.

Random Domain Insertion. We have previously described a method for creating molecular switches for which JHU (JHU ref 1706) has filed a provisional patent (60/362,588) and a PCT application. A schematic representation of an exemplary implementation of this method is shown in Figure 1. The method creates a library of random insertions of one gene (or gene fragment) into another gene (or gene fragment). From this library molecular switches can be selected or screened for.

We demonstrated this method by creating two allosteric enzymes (i.e. enzymatic molecular switches) (Guntas and Ostermeier 2004). These allosteric enzymes have been created by the covalent linkage of non-interacting, monomeric proteins with the prerequisite effector-binding and catalytic functionalities, respectively. The fragment of the TEM-1 β-lactamase gene coding for the mature protein lacking its signal sequence was randomly inserted into the E. coli maltose binding protein (MBP) gene to create a domain insertion library. This library's diversity derived both from the site of insertion and from a distribution of tandem duplications or deletions of a portion of the MBP gene at the insertion site. From a library of ~$2x10^4$ in-frame fusions, ~800 library members conferred a phenotype to E. coli cells that was consistent with the presence of

bifunctional fusions that could hydrolyze ampicillin and transport maltose in E. coli. Partial screening of this bifunctional sublibrary resulted in the identification of two enzymes in which the presence of maltose modulated the rate of nitrocefin hydrolysis. For one of these enzymes, the presence of maltose increased k_{cat} by 70% and k_{cat}/K_m by 80% and resulted in kinetic parameters that were almost identical to TEM-1 β-lactamase. Such an increase in activity was only observed with maltooligosaccharides whose binding to MBP is known to induce a conformational change from the open to the closed form. Modulation of the rate of nitrocefin hydrolysis could be detected at maltose concentrations less than 1 μM. Intrinsic protein fluorescence studies were consistent with a conformational change being responsible for the modulation of activity.

Other strategies for creating molecular switches. There are five predominant existing strategies for creating protein molecular switches: (1) domain insertion(discussed above), (2) control of oligomerization or proximity using chemical inducers of dimerization (CID), (3) chemical rescue, (4) fusion of the target protein to a steroid binding domain (SBD) and (5) coupling proteins to nonbiological materials such as 'smart' polymers (Stayton, Shimoboji et al. 1995; Ding, Fong et al. 2001; Kyriakides, Cheung et al. 2002) or metal nanocrystals (Hamad-Schifferli, Schwartz et al. 2002). A sixth related strategy, control by regulation of aggregation (Rivera, Wang et al. 2000), is, strictly speaking, is a method for controlling a protein's availability and not a soluble protein's activity. A seventh strategy, introducing proximal cysteine residues into a protein and then modulating a protein's activity by controlling whether the cysteine forms a disulfide is obviously very limited in the signals that can be employed (ref).

Control using a chemical inducer of dimerization (CID). This strategy utilizes a synthetic ligand as the CID that controls the oligomeric or proximity of two proteins (Spencer, Wandless et al. 1993), best exemplified by signal transduction and gene expression (Rivera 1998). The CIDs are small molecules that have two binding surfaces that facilitate the dimerization of domains fused to target proteins. This was first developed using the immunosupressant FK506 to facilitate dimerization of target proteins fused to the FK506-binding protein, FKBP12 (Spencer, Wandless et al. 1993). Several variations on this system as well as a system using the antibiotic coumermycin to dimerize proteins fused to B subunit of bacterial DNA gyrase (GyrB) (Farrar, Olson et al. 2000) have appear since. CIDS have been used to initiate signaling pathways by dimerizing receptors on the cell surface, to translocate cytosolic proteins to the plasma membrane, to import and export proteins from the nucleus, to induce apoptosis and to regulate gene transcription (Bishop, Buzko et al. 2000; Farrar, Olson et al. 2000). However, CIDs have only been applied to those functions that require changes in the oligomeric state or proximity of the two proteins. As described in the literature, this approach can not be readily applied to a single protein.

Chemical Rescue. Chemical rescue is the restoration of activity to a mutant, catalytically defective enzyme by the introduction of a small molecule that has the requisite properties of the mutated residues. Since first described for subtilisin (Carter and Wells 1987), chemical rescue has been demonstrated for a number of different mutated protein-small molecule pairs (Williams, Wang et al. 2000). The vast majorities of these rescues required > 5 mM to showed detectable rescue and the maximum fold improvement in activity of the mutant was generally less than 100-fold and required >100

mM concentrations of the rescuing molecule. Chemical rescue has only recently been applied as a strategy for control, in this case of dimerization (Guo, Zhou et al. 2000).

Fusion to a steroid binding domain. The protein to be controlled is fused end-to-end to a SBD (Picard 2000). In the absence of the steroid that binds to the SBD, it is believed that a Hsp90-SBD complex sterically interferes with the activity of the protein fused to the SBD. The disassembly of the complex upon steroid binding restores activity to the protein. This strategy has been successfully applied, principally to transcription factors and kinases (Picard 2000). Artificial transcription factors using this strategy have been developed (called GeneSwitch) and have promise for tissue-specific gene expression in transgenic animals and human gene therapy (Burcin, BW et al. 1998; Burcin, Schiedner et al. 1999).

Coupling to non-biological materials The protein to be controlled is coupled to a non-biological material that responds to an external and thereby affects the protein coupled to it. 'Smart' polymers that change their conformation upon a change in pH or temperature have been conjugated to proteins near ligand-binding sites to create switches that sterically block access to the binding site at, for example higher temperature but not at lower temperatures (Stayton, Shimoboji et al. 1995; Ding, Fong et al. 2001). Inductive coupling of a magnetic field to metal nanocrystals attached to biomolecules resulting in an increase in local temperature thereby inducing denaturation, has so far only been applied to DNA (Hamad-Schifferli, Schwartz et al. 2002) but presumably will also work with proteins.

Limitations of existing approaches. Domain insertion is the most promising and most general strategy for engineering a molecular switch, as the other strategies are limited in the sorts of signals that can be employed or the types of proteins that can be controlled. CIDs have only been applied to those functions that require changes in the oligomeric state or proximity of the two proteins and thus cannot be used to control a single protein. The serious limitation of the chemical rescue approach is the inability to apply the method to any signal and the lack of sensitivity (high concentrations of the signal are required for a small change in activity). There are several factors that limit the SBD strategy as a general method for controlling any protein. End-to-end fusion appears applicable only to fusions with SBDs as no other similar system has appeared in the thirteen years since its introduction (Picard, Salser et al. 1988; Eilers, Picard et al. 1989). Thus, the potential for the creation of molecular switches triggered by non-steroid signals by this method appears very limited. Also, a steroid is used to impart control, but steroids have pleiotropic effects and thus are not likely to be useful for engineering therapeutic molecules. Finally, the system requires a fourth component, the Hsp90 complex and apparently has low induction ratios (about ten fold) (Spencer 1996). Coupling to non-biological systems are limited in the signals that can be employed. Primarily the only signals that have been employed are pH and temperature and it is difficult to image that this strategy could be used for signals that are proteins or metabolites.

There are several limitations of existing domain insertion strategies: (1) With the exception of our work, all domain insertion studies for creating molecular switches have examined a very small number of possible insertional fusions between the two domains (i.e. the insertion locations were rationally chosen). Just as combinatorial approaches have proven to be invaluable for improving protein function by directed evolution

(Arnold 2001) (e.g. random mutagenesis and DNA shuffling), the development of combinatorial methods for domain insertion will likely have a similar effect on the development of molecular switches. (2) The switching behavior achieved my existing methods was generally modest (less than 2-fold effect). A combinatorial approach systematically tests a plurality of potential switches to find the optimum switch. (3) None of the studies have explored circular permutation of either the target or the inserted gene; thus, existing domain insertion strategies can only explore a limited number of geometric configurations between the two domains.

Circular permutation. A circularly permuted protein has its original N- and C-termini fused and new N- and C- termini created by a break elsewhere in the sequence. The first *in vitro* construction of a circular permuted protein was carried out on bovine pancreatic trypsin inhibitor by chemical means (Goldenberg and Creighton 1983). Since then, a number of circular permuted proteins have been constructed (Heinemann and Hahn 1995), primarily by genetic methods, including TEM-1 β-lactamase (Osuna, Pérez-Blancas et al. 2002) used in this study. These studies have shown that circular permuted proteins very often fold up into stable, active proteins. Comparisons of primary and tertiary structures within several protein families have led to the conclusion that circular permutation occurs in natural protein sequences.

Random circular permutation. A genetic method for *random* circular permutation of any gene was first described by Graf and Schachmann (Graf and Schachman 1996). The method is comprised of the following steps: (i) isolation of a linear fragment of double stranded DNA (of the gene to be randomly circularly permuted) with flanking compatible ends, (ii) cyclization of this DNA fragment by ligation under dilute conditions, (iii) random linearization of the cyclized gene using DNase I digestion in the presence of Mn^{2+} at dilute concentrations of the enzyme such that the DNase I, on average, makes one double strand break, (iv) repair of nicks and gaps using a DNA polymerase and a DNA ligase, and (v) ligation of the fragment into a desired vector by blunt end ligation to create the plasmid library of randomly circularly permuted genes. The procedure has been used to systematically identify permissive sites for circular permutation in aspartate transcarbamoylase (Graf and Schachman 1996), DsbA (Hennecke, Sebbel et al. 1999) and GFP (Baird, Zacharias et al. 1999) and DHFR (Iwakura, Nakamura et al. 2000) and in most case a variety of permissible sites were identified. The original random circular permutation protocol (ref) as published works poorly, if at all and likely had errors in it (Graf and Schachman 1996). The methodology for randomly circularly permuting DNA was improved upon in a study involving the random localization of a restriction enzyme site in a fixed length of DNA (Ostermeier and Benkovic 2001).

Novelty. To our knowledge, a circularly permuted gene that has been circularly permuted at a specific site and inserted into another gene (i.e. by domain insertion at a specific site or by random domain insertion) has never been described, though such a construct has been proposed (but not tested) as having potential for creating GFP sensors (Baird, Zacharias et al. 1999). To our knowledge, the insertion of a randomly circularly permuted gene into another gene (i.e. by domain insertion at a specific site or by random domain insertion) has never been described. Furthermore, the idea that such a process will create a molecular switch does not have precedent and is not obvious.

Description of the Invention

This invention combines circular permutation and domain insertion in a novel fashion to create molecular switches. The invention is reduced to practice using TEM-1 β-lactamase (BLA) and *E. coli* maltose binding protein (MBP) as model proteins. These are the same proteins used to demonstrate how random domain insertion alone can create switches (our previous invention). Although the random domain insertion method alone has been able to create a BLA-MBP molecular switch, to date the magnitude of the switching achieved has been modest (< 2-fold) and only "on-switches" were found (i.e. the addition of signal resulted in an increase in function). The switching obtained using circular permutation and domain insertion has resulted in both "on" and "off" switches and magnitude of the switching has been much higher (in one case, at least 32-fold). Although demonstrated with two particular proteins this invention applies to functionally coupling any two proteins.

Key to this invention is the circular permutation of one of the genes (in this case the insert gene, although circular permutation of the acceptor sequence is also possible). A β-lactamase that has been circularly permuted and inserted into MBP will have a different covalent linkage and a different spatial orientation to the MBP compared to the original insertion using the wildtype N- and C-termini of BLA. This potentially can have dramatic effects on the switching behavior. For example, the switching behavior can change due to (1) decreasing the distance from the active site of β-lactamase to residues allosterically linked to maltose binding to MBP, (2) exploring new paths by which the effects of maltose binding can propagate to the active site of β-lactamase and (3) changing the stability of the BLA domain making it more susceptible to changes in conformation of the MBP domain. For example, the switches identified by domain insertion of a non-circularly permuted BLA (Guntas and Ostermeier 2004) were linked through a very stable portion of the β-lactamase (Luque and Freire 2000). Circularly permuting β-lactamase will allow linkage through less stable regions that are postulated to be more sensitive to the effects of maltose binding.

Two strategies for creating molecular switches involving random circular permutation of the insert gene have been demonstrated (Figure 2) using MBP and BLA as an example. In the first, (called "Random Circular Permutation of Insert and Domain Insertion at a Specific Site") the BLA gene is circularly permuted and inserted into a specific site in the MBP. This site could be a site previously shown to be useful for creating molecular switches (as is demonstrated here) or a site that is predicted, by computational methods or other means, to be useful in creating a molecular. In the second method (called "Random Circular Permutation of Insert and Random Domain Insertion") the BLA gene is randomly circularly permuted and randomly inserted into a plasmid containing the MBP gene. The use of BLA and MBP are only examples and potentially any two proteins can be functionally coupled in this manner.

Linkers for Circular Permutation. In order to circularly permute a gene it is usually necessary to include DNA that codes for a linker to link the original N- and C-termini. We chose to test two different linkers. In the first (the "DKS linker") the β-lactamase was randomly circularly permuted by fusing the 5'- and 3'- ends with a DNA sequence coding for the tripeptide linker DKS previously found in a combinatorial library of linkers to be most conducive for circularly permuting β-lactamase when the new N- and C-termini were located at a specific location (Osuna, Pérez-Blancas et al. 2002). In

the second (the "GSGGG linker") the β-lactamase was randomly circularly permuted by fusing the 5'- and 3'- ends with a DNA sequence coding for the flexible pentapeptide linker GSGGG.

Preparation of BLA Insert DNA. The β-lactamase gene fragment *bla [24-286]* (codes for amino acids 24-286) was amplified by PCR from pBR322 such that it was flanked by *Ear*I or *BamH*I sites and sequences coding for the linkers described above and cloned into pGem T-vector (Promega) to create pBLA-CP(DKS) and pBLA-CP(GSGGG), respectively (Figures 3 and 4). DNA coding for amino acids 1-23 were not desired because it codes for the signal sequence that targets ß-lactamase to the periplasm. This sequence gets cleaved upon entering the periplasm and is not part of the mature, active ß-lactamase.

One hundred and thirty micrograms of pBLA-CP(GSGGG) was digested with 2000 units of *BamH*I and 140 micrograms of pBLA-CP(DKS) was digested with 600 units of *Ear*I in the buffers and conditions recommended by the manufacturer of the restriction enzyme. The fragment containing the BLA gene was purified by agarose gel electrophoresis using the QIAquick gel purification kit. This DNA was treated with T4 DNA ligase under dilute concentrations to cyclize the DNA (18 hours at 16 °C with 600 Weiss units of T4 DNA Ligase in the presence of 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 ug/ml BSA in a total volume of 5.1 ml). The ligation reaction was stopped by incubation at 65 °C for 20 minutes. The DNA was concentrated by vacufuge and desalted using the QIAquick PCR purification kit. The circular fragment was purified by agarose gel electrophoresis using the QIAquick gel purification kit.

The conditions for DNaseI digestion were determined experimentally by adding different amounts of DNaseI and analyzing the digested products by agarose gel electrophoresis. The digestion conditions were chosen such that a significant fraction of DNA was undigested in order maximize the amount of linear DNA that only had one double strand break. In general, approximately 1 milliunit of DNaseI per microgram of DNA (at a concentration 10 micrograms/ml) of for an 8 minute digestion at 22 °C is close to being optimum. Sometimes more or less DNaseI was required and thus it is recommended that for each library constructed the correct amount of DNaseI be determined experimentally by test digestions. The following conditions are given as a representative example. Six micrograms of circular DNA was digested with 6 milliunits of DNase I (Roche) for 8 minutes at 22 °C in the presence of 50 mM TrisHCl (pH 7.4), 1 mM $MnCl_2$ and 50 micrograms/ml BSA in 0.6 ml reaction volume. The reaction was stopped by adding EDTA to a concentration of 5 mM. The DNA was desalted using the QIAquick PCR purification kit and repaired by 6 units of T4 DNA polymerase and 6 Weiss Units of T4 DNA ligase at 12 °C for 15 minutes in the presence of 100 micromolar dNTP, 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP and 25 ug/ml BSA. The repaired, linear DNA was purified by agarose gel electrophoresis using the QIAquick gel purification kit. This is the circularly permuted gene that is ready for insert into another plasmid.

Preparation of target DNA for random domain insertion libraries. Forty μg of pDIM-C8-Mal was digested with DNaseI (0.01 units) for 8 minutes at 22°C in the presence of 50 mM Tris-HCl, pH 7.4, 10 mM $MnCl_2$ and 50 μg/ml BSA in a total volume of 1 ml. The reaction was quenched by the addition of EDTA to a concentration of 5 mM

and the solution was desalted using four Qiaquick PCR purification columns into 200 µl elution buffer which was subsequently concentrated by vacufuge. Nicks and gaps were repaired by incubating at 12°C for 1 hour in a total volume of 120 µl in the presence of T4 DNA polymerase (15 units) and T4 DNA ligase (12 Weiss units) in the presence of 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 µg/ml BSA and 125 µM dNTPs. The reaction was stopped by incubating at 80°C for 10 minutes. Sodium chloride was added to 100 mM and the DNA was dephosphorylated by adding alkaline phosphatase (60 units) and incubating at 37°C for 1 hour. The DNA was desalted as before and the linear DNA (corresponding to the randomly linearized pDIM-C8-Mal) was isolated from circular forms of the plasmid by agarose gel electrophoresis using the Qiaquick gel purification kit.

Preparation of target DNA for site-specific insertion libraries. Plasmid pDIM-C8-Mal was modified using overlap extension (Horton, Hunt et al. 1989) to be suitable for insertion of the circularly permuted BLA at two specific sites: (a) between MBP [1-165] and MBP [164-370] and (b) at the C-terminus of MBP. The plasmids were modified in analogous ways, the modifications for insertion between MBP [1-165] and MBP [164-370] to create plasmid pDIMC8-MBP(164-165) is described as an example (Figure 5). Two inverted *Sap*I sites were inserted between DNA coding for MBP [1-165] and MBP [164-370] in such a manner that digestion with *Sap*I and subsequent filling in of the resulting overhangs using Klenow polymerase in the presence of dNTPs results in a perfectly blunt *MBP [1-165]* on one side and a perfectly blunt *MBP [164-370]* on the other. This is achieved by virtue of the fact that *Sap*I is a type IIS restriction enzyme that cuts outside of its recognition sequence. Other type IIS restriction enzymes could have been used. Non-type IIS restriction enzymes could also be used if it is acceptable to have their recognition site as part of the gene fragment that is being inserted into.

Three micrograms of pDIMC8-MBP(164-165) was digested with 6 units of *Sap*I at 37 °C in the presence of 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol (pH 7.9), 100 ug/ml BSA for 2.5 hours. The DNA was desalted using the QIAquick PCR purification kit and repaired with 5 units of Klenow at 25 °C for 20 minutes in the presence of 33 micromolar dNTPs, 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$ and 1 mM dithiothreitol (pH 7.9). The enzyme was heat inactivated by incubation at 75 °C for 20 minutes. Sodium chloride was added to 100 mM and ten units of Calf Intestinal Phosphatase was added and the solution was incubated for 1 hour at 37 °C. Dephosphorylation was performed to prevent recircularization of the vector without receiving an insert in the subsequent ligation step. The vector DNA was purified by agarose gel electrophoresis using the QIAquick gel purification kit.

Ligation of inserts into target DNA. Insert DNA (85 ng) comprising the circularly permuted BLA was ligated to the prepared target DNA (100 ng) at 22°C overnight in the presence of T4 DNA ligase (30 Weiss units) and the ligase buffer provided by the manufacturer in a total volume of 13 µl. After ethanol precipitation, 10% of the ligase-treated DNA was electroporated into 50 µl Electromax DH5α-E electrocompetent cells. Transformed cells were plated on large (248 mm x 248 mm) LB agar plate supplemented with 50 µg/ml chloramphenicol (Cm). The naïve domain insertion library was recovered from the large plate (Ostermeier, Nixon et al. 1999) and stored in frozen aliquots.

Screening for allosteric enzymes. The libraries were diluted from frozen aliquots and plated on LB plates containing different concentrations of ampicillin (Tables 1 and 2). A number of colonies were picked (Table 4) and grown in LB overnight in 96 well plates (0.5 ml/well) in the presence of 1 mM IPTG and 50 μg/ml Cm. Next, 50 μl of PopCulture (Novagen) and 2.5 unit of benzonase nuclease was added to each well and incubated for 15 minutes at room temperature to lyse the cells. The cells debris and any unlysed cells were pelleted by centrifugation and supernatant was recovered. In 96-well format, 60 μl of lysate was assayed for hydrolysis of nitrocefin (50 μM) by monitoring the increase in absorbance at 490 nm in 100 mM sodium phosphate buffer, pH 7.0, both with and without 5 mM maltose. Any lysate in which there was a difference in rate of more than 2-fold (between with and without maltose) was selected for retesting and further investigation.

Confirmation and identification of positives. Library members identified as having more than 200% switching activity in the 96-well plate screen were grown 24-48 hours in 100 ml LB media in 500 ml shake flasks at 22°C without IPTG. The cells were pelleted and resuspended in 8 ml assay buffer (100 mM sodium phosphate buffer, pH 7.0) and lysed by French press. The soluble fraction of this lysate was assayed for hydrolysis of nitrocefin (50 μM) at 22 °C as previously described (Guntas and Ostermeier 2004) both with and without 5 mM maltose. Initial rates were determined from absorbance at 486 nm monitored as a function of time. The enzyme was incubated at the assay temperature in the absence or presence of 5 mM maltose for four minutes prior to performing the assay. All assays contained 100 mM sodium phosphate buffer, pH 7.0. Library members for which there was a difference in the initial rate of more than about 2-fold were sequenced (Table 4). Switches RG-5-169 and RG-200-13 were also assayed in the presence of 5 mM sucrose or 5 mM glucose. Neither sugar affected the velocity of nitrocefin hydrolysis indicating that the switching effect was specific for maltose, a ligand to which MBP binds.

Analysis of Purified Switch RG-200-13. A 6xHis tag was added to the C-terminus of RG-200-13 and the fusion was purified as previously described for another switch (Guntas and Ostermeier 2004). The protein was purified to approximately 60% purity. The kinetic constants and binding constants were determined from Eadie-Hofstee plots and Eadie plot equivalents, respectively, using a spectrophotometric assay for nitrocefin hydrolysis. Initial rates for nitrocefin hydrolysis were determined from absorbance at 486 nm monitored as a function of time. The enzyme was incubated at the assay temperature in the absence or presence of saccharide for four minutes prior to performing the assay. All assays contained 100 mM sodium phosphate buffer, pH 7.0. The dissociation constant for maltose was determined using change in velocity of nitrocefin hydrolysis as a signal.

Only sugars known to bind to MBP had an effect on nitrocefin hydrolysis (Table 5). Those sugars that produce a large conformational change upon binding MBP (Quiocho, Spurlino et al. 1997) (maltose and maltotriose) produced the largest change in the velocity of nitrocefin hydrolysis. Beta-cyclodextrin, which produces a small conformational change upon binding MBP (Evenas, Tugarinov et al. 2001), has a small effect. The effect of maltotetraitol is intermediate, consistent with the fact that maltotetraitol-binding to MBP results in a mixture of open and closed structures (Duan, Hall et al. 2001).

The kinetic parameters of RG-200-13 are reported in Table 6. The kinetic parameters of RG-200-13 at 22°C in the presence of maltose (k_{cat} = ~520 s^{-1}; K_m = ~85 μM) are very similar to previously reported values for TEM-1 β-lactamase at 30 °C (k_{cat} = 930 s^{-1}; K_m = 52 μM) (Raquet, Lamotte-Brasseur et al. 1994) indicating that RG-200-13 is essentially a fully functional TEM-1 β-lactamase in the presence of maltose. The k_{cat}/K_m in the presence of 5 mM maltose is approximately 25-fold higher than in the absence of maltose. The K_d for maltose binding to RG-200-13 at 22°C was ~5 μM, similar to the K_d previously reported for maltose binding to MBP (1-1.5 μM) (Schwartz, Kellermann et al. 1976).

The effect of 5 mM maltose on other sustrates of BLA is shown in Table 7. Maltose binding had the largest effect on cephalothin (of the substrates tested), with the velocity of cephalothin hydrolysis being 32-fold higher in the presence of maltose than in its absence. Based on the effects on other substrates, the actual switching effect on k_{cat}/K_m for cephalothin is likely to be much higher than 32-fold.

The fact that the magnitude of the switching effect of RG-200-13 is substrate identity and concentration dependent strongly argues that maltose is converting the protein from a less active to a more active conformation. The alternative explanation that maltose affects the equilibrium between unfolded (inactive) and folded (active) forms of the protein would result in a switching effect that was independent of the substrate being tested (and independent of substrate concentration), which is not the case.

Applications of the molecular switches obtained by the invention. Protein molecular switches are proteins whose activity can be modulated through a signal such as the binding of a small molecule, the interaction with another protein or the sensing of some other signal (e.g. a change in pH). In other words, a molecular switch functionally couples an external signal to functionality. Protein molecular switches engineered by the methods of this invention have a wide variety of potential applications (Fig 6) in, for example, (a) the regulation of gene transcription, (b) the modulation of cell signaling pathways, (c) targeted drug delivery, (d) drug transport, (e) the creation of conditionally active toxic proteins, (f) metabolic engineering, and (g) the creation of novel biosensors.

The BLA-MBP switches have potential as in vitro or in vivo molecular sensors. BLA activity can be measure *in vivo* or *in vitro* using specially designed fluorescent substrate (Zlokarnik, Negulescu et al. 1998; Gao, Xing et al. 2003). MBP is a member of a class of proteins called periplasmic binding proteins that have been shown to be amenable to computational design to alter the proteins so that they bind new ligands (Looger, Dwyer et al. 2003). Thus, the BLA-MBP switches potentially can be used to detect different ligands *in vivo* as enzymatic activity would be proportional to ligand concentration.

References

Geneswitch, http://www.geneswitch.com/. **2002**.

Arnold, F. H. (2001). "Combinatorial and computational challenges for biocatalyst design." Nature **409**(6817): 253-7.

Ataka, K. and V. A. Pieribone (2002). "A genetically targetable fluorescent probe of channel gating with rapid kinetics." Biophys J **82**(1 Pt 1): 509-16.

Aÿ, J., F. Götz, et al. (1998). "Structure and function of the *Bacillus* hybrid enzyme GluXyn-1: native-like jellyroll fold preserved after insertion of autonomous globular domain." Proc.Natl.Acad.Sci.USA **95**: 6613-6618.

Baird, G. S., D. A. Zacharias, et al. (1999). "Circular permutation and receptor insertion within green fluorescent proteins." Proc.Natl.Acad.Sci.U.S.A **96**(20): 11241-11246.

Barany, F. (1985). "Single-stranded hexameric linkers: a system for in-phase insertion mutagenesis and protein engineering." Gene **37**(1-3): 111-23.

Barany, F. (1985). "Two-codon insertion mutagenesis of plasmid genes by using single-stranded hexameric oligonucleotides." Proc Natl Acad Sci U S A **82**(12): 4202-6.

Barany, F. (1987). "A genetic system for isolation and characterization of TaqI restriction endonuclease mutants." Gene **56**(1): 13-27.

Betton, J.-M., J. P. Jacob, et al. (1997). "Creating a bifunctional protein by insertion of β-lactamase into the maltodextrin-binding protein." Nat.Biotechnology **15**: 1276-1279.

Bibi, E. and O. Beja (1994). "Membrane topology of multidrug resistance protein expressed in Escherichia coli. N-terminal domain." J Biol Chem **269**(31): 19910-5.

Biondi, R. M., P. J. Baehler, et al. (1998). "Random insertion of GFP into the cAMP-dependent protein kinase regulatory subunit from *Dictyostelium discoideum*." Nucleic Acids Res. **26**(21): 4946-4952.

Bishop, A., O. Buzko, et al. (2000). "Unnatural ligands for engineered proteins: new tools for chemical genetics." Annu Rev Biophys Biomol Struct **29**: 577-606.

Boeke, J. D. (1981). "One and two codon insertion mutants of bacteriophage f1." Mol Gen Genet **181**(3): 288-91.

Burcin, M. M., O. M. BW, et al. (1998). "A regulatory system for target gene expression." Front Biosci **3**: c1-7.

Burcin, M. M., G. Schiedner, et al. (1999). "Adenovirus-mediated regulable target gene expression in vivo." Proc Natl Acad Sci U S A **96**(2): 355-60.

Carter, P. and J. A. Wells (1987). "Engineering enzyme specificity by "substrate-assisted catalysis"." Science **237**(4813): 394-9.

Collinet, B., M. Herve, et al. (2000). "Functionally accepted insertions of proteins within protein domains." J Biol Chem **275**(23): 17428-33.

Cosgriff, A. J. and A. J. Pittard (1997). "A topological model for the general aromatic amino acid permease, AroP, of Escherichia coli." J Bacteriol **179**(10): 3317-23.

de Lorimier, R. M., J. J. Smith, et al. (2002). "Construction of a fluorescent biosensor family." Protein Sci **11**(11): 2655-75.

Ding, Z., R. B. Fong, et al. (2001). "Size-dependent control of the binding of biotinylated proteins to streptavidin using a polymer shield." Nature **411**(6833): 59-62.

Doi, N., M. Itaya, et al. (1997). "Insertion of foreign random sequences of 120 amino acid residues into an active enzyme." FEBS Lett. **402**(2-3): 177-180.

Doi, N. and H. Yanagawa (1999). "Design of generic biosensors based on green fluorescent proteins with allosteric sites by directed evolution." FEBS Lett. **453**: 305-307.

Duan, X., J. A. Hall, et al. (2001). "Crystal structures of the maltodextrin/maltose-binding protein complexed with reduced oligosaccharides: flexibility of tertiary structure and ligand binding." J Mol Biol **306**(5): 1115-26.

Ehrmann, M., D. Boyd, et al. (1990). "Genetic analysis of membrane protein topology by a sandwich gene fusion approach." Proc Natl Acad Sci U S A **87**(19): 7574-8.

Eilers, M., D. Picard, et al. (1989). "Chimaeras of myc oncoprotein and steroid receptors cause hormone- dependent transformation of cells." Nature **340**(6228): 66-8.

Evenas, J., V. Tugarinov, et al. (2001). "Ligand-induced structural changes to maltodextrin-binding protein as studied by solution NMR spectroscopy." J Mol Biol **309**(4): 961-74.

Farrar, M. A., S. H. Olson, et al. (2000). "Coumermycin-induced dimerization of GyrB-containing fusion proteins." Methods Enzymol **327**: 421-9.

Fehr, M., W. B. Frommer, et al. (2002). "Visualization of maltose uptake in living yeast cells by fluorescent nanosensors." Proc Natl Acad Sci U S A **99**(15): 9846-51.

Freimuth, P. I. and H. S. Ginsberg (1986). "Codon insertion mutants of the adenovirus terminal protein." Proc Natl Acad Sci U S A **83**(20): 7816-20.

Gao, W., B. Xing, et al. (2003). "Novel fluorogenic substrates for imaging B-Lactamase Gene Expression." J.Am.Chem.Soc. **125**: 11146-11147.

Goldenberg, D. P. and T. E. Creighton (1983). "Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor." J.Mol.Biol **165**(2): 407-413.

Graf, R. and H. K. Schachman (1996). "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase." Proc.Natl.Acad.Sci.USA **93**: 11591-11596.

Guerrero, G. and E. Y. Isacoff (2001). "Genetically encoded optical sensors of neuronal activity and cellular function." Curr Opin Neurobiol **11**(5): 601-7.

Guntas, G. and M. Ostermeier (2004). "Creation of an allosteric enzyme by domain insertion." J Mol Biol **in press**.

Guo, Z., D. Zhou, et al. (2000). "Designing small-molecule switches for protein-protein interactions." Science **288**(5473): 2042-5.

Hallet, B., D. J. Sherratt, et al. (1997). "Pentapeptide scanning mutagenesis: random insertion of a variable five amino acid cassette in a target protein." Nucleic Acids Res. **25**(9): 1866-1867.

Hamad-Schifferli, K., J. J. Schwartz, et al. (2002). "Remote electronic control of DNA hybridization through inductive coupling to an attached metal nanocrystal antenna." Nature **415**(6868): 152-5.

Heinemann, U. and M. Hahn (1995). "Circular permutation of polypeptide chains: implications for protein folding and stability." Prog.Biophys.Mol.Biol **64**(2-3): 121-143.

Hennecke, J., P. Sebbel, et al. (1999). "Random circular permutation of DsbA reveals segments that are essential for protein folding and stability." J.Mol.Biol **286**(4): 1197-1215.

Horton, R. M., H. D. Hunt, et al. (1989). "Engineered hybrid genes without the use of restriction enzymes: gene splicing by overlap extension." Gene **77**: 61-68.

Iwakura, M., T. Nakamura, et al. (2000). "Systematic circular permutation of an entire protein reveals essential folding elements." Nat Struct Biol **7**(7): 580-5.

Jones, S., M. Stewart, et al. (1998). "Domain assignment for protein structures using a consensus approach: characterization and analysis." Protein Sci **7**(2): 233-42.

Kratz, P. A., B. Bottcher, et al. (1999). "Native display of complete foreign protein domains on the surface of hepatitis B virus capsids." Proc Natl Acad Sci U S A **96**(5): 1915-20.

Kyriakides, T. R., C. Y. Cheung, et al. (2002). "pH-sensitive polymers that enhance intracellular drug delivery in vivo." J Control Release **78**(1-3): 295-303.

Lacatena, R. M., A. Cellini, et al. (1994). "Topological analysis of the human beta 2-adrenergic receptor expressed in Escherichia coli." Proc Natl Acad Sci U S A **91**(22): 10521-5.

Ladant, D., P. Glaser, et al. (1992). "Insertional mutagenesis of Bordetella pertussis adenylate cyclase." J.Biol Chem. **267**(4): 2244-2250.

Looger, L. L., M. A. Dwyer, et al. (2003). "Computational design of receptor and sensor proteins with novel functions." Nature **423**(6936): 185-90.

Lu, Z., K. S. Murray, et al. (1995). "Expression of thioredoxin random peptide libraries on the Escherichia coli cell surface as functional fusions to flagellin: a system designed for exploring protein-protein interactions." Biotechnology (N Y) **13**(4): 366-72.

Luque, I. and E. Freire (2000). "Structural stability of binding sites: consequences for binding affinity and allosteric effects." Proteins **Suppl**(4): 63-71.

Nagai, T., A. Sawano, et al. (2001). "Circularly permuted green fluorescent proteins engineered to sense Ca2+." Proc Natl Acad Sci U S A **98**(6): 3197-202.

Ostermeier, M. and S. J. Benkovic (2001). "Construction of hybrid gene libraries involving the circular permutation of DNA." Biotechnol. Lett. **23**: 303-310.

Ostermeier, M., A. E. Nixon, et al. (1999). "Combinatorial protein engineering by incremental truncation." Proc.Natl.Acad.Sci.USA **96**: 3562-3567.

Osuna, J., A. Pérez-Blancas, et al. (2002). "Improving a circularly permuted TEM-1 ß-lactmase by directed evolution." Protein Eng **15**(6): 463-470.

Pi, J. and A. J. Pittard (1996). "Topology of the phenylalanine-specific permease of Escherichia coli." J Bacteriol **178**(9): 2650-5.

Picard, D. (2000). "Posttranslational regulation of proteins by fusions to steroid-binding domains." Methods Enzymol **327**: 385-401.

Picard, D., S. J. Salser, et al. (1988). "A movable and regulable inactivation function within the steroid binding domain of the glucocorticoid receptor." Cell **54**(7): 1073-80.

Pigeon, R. P. and R. P. Silver (1994). "Topological and mutational analysis of KpsM, the hydrophobic component of the ABC-transporter involved in the export of polysialic acid in Escherichia coli K1." Mol Microbiol **14**(5): 871-81.

Quiocho, F. A., J. C. Spurlino, et al. (1997). "Extensive features of tight oligosaccharide binding revealed in high- resolution structures of the maltodextrin transport/chemosensory receptor." Structure **5**(8): 997-1015.

Raquet, X., J. Lamotte-Brasseur, et al. (1994). "TEM beta-lactamase mutants hydrolysing third-generation cephalosporins. A kinetic and molecular modelling analysis." J Mol Biol **244**(5): 625-39.

Rivera, V. M. (1998). "Controlling gene expression using synthetic ligands." Methods **14**(4): 421-9.

Rivera, V. M., X. Wang, et al. (2000). "Regulation of protein secretion through controlled aggregation in the endoplasmic reticulum." Science **287**(5454): 826-30.

Russell, R. B. and C. P. Ponting (1998). "Protein fold irregularities that hinder sequence analysis." Curr.Opin.Struct.Biol **8**(3): 364-371.

Sarsero, J. P. and A. J. Pittard (1995). "Membrane topology analysis of Escherichia coli K-12 Mtr permease by alkaline phosphatase and beta-galactosidase fusions." J Bacteriol **177**(2): 297-306.

Schwartz, M., O. Kellermann, et al. (1976). "Further studies on the binding of maltose to the maltose-binding protein of Escherichia coli." Eur J Biochem **71**(1): 167-70.

Siegel, M. S. and E. Y. Isacoff (1997). "A genetically encoded optical probe of membrane voltage." Neuron **19**(4): 735-41.

Siegel, M. S. and E. Y. Isacoff (2000). "Green fluorescent protein-based sensors for detecting signal transduction and monitoring ion channel function." Methods Enzymol **327**: 249-59.

Spencer, D. M. (1996). "Creating conditional mutations in mammals." Trends Genet **12**(5): 181-7.

Spencer, D. M., T. J. Wandless, et al. (1993). "Controlling signal transduction with synthetic ligands." Science **262**(5136): 1019-24.

Starzyk, R. M., J. J. Burbaum, et al. (1989). "Insertion of new sequences into the catalytic domain of an enzyme." Biochemistry **28**(21): 8479-8484.

Stayton, P. S., T. Shimoboji, et al. (1995). "Control of protein-ligand recognition using a stimuli-responsive polymer." Nature **378**(6556): 472-4.

Stone, J. C., T. Atkinson, et al. (1984). "Identification of functional regions in the transforming protein of Fujinami sarcoma virus by in-phase insertion mutagenesis." Cell **37**(2): 549-58.

Tucker, C. L. and S. Fields (2001). "A yeast sensor of ligand binding." Nat Biotechnol **19**(11): 1042-6.

Williams, D. M., D. Wang, et al. (2000). "Chemical rescue of a mutant protein-tyrosine kinase." J Biol Chem **275**(49): 38127-30.

Zebala, J. and F. Barany (1991). "Mapping catalytically important regions of an enzyme using two-codon insertion mutagenesis: a case study correlating β-lactamase mutants with the three-dimensional structure." Gene **100**: 51-57.

Zlokarnik, G., P. A. Negulescu, et al. (1998). "Quantitation of transcription and clonal selection of single living cells with beta-lactamase as reporter." Science **279**(5347): 84-8.

JD

Table 1 Library statistics

Insertion site in MBP	Linker in BLA	Library size (Number of transformants with BLA insert).	Number of library members that can grow on 50 μg/ml AMP (see Table 2)	Number of colonies screened for switching (see Table 3)	Number of unique switches found with ≥ 2-fold effect*	Increase in velocity (of nitrocefin hydrolysis in presence of maltose) of best switch
164-165	DKS	0.44×10^6	515	848	2	+97%
	GSGGG	1.05×10^6	361	1248	1	-250%
C-terminus	DKS	1.03×10^6	2414	576	0	
	GSGGG	0.30×10^6	1615	1920	1-4	+234%
random	DKS	0.41×10^6	191	384	0	
	GSGGG	1.20×10^6	1156	3312	5	+1650%

* ≥ 2-fold change in velocity of nitrocefin hydrolysis in the presence of 5 mM maltose.

Table 2 Number of Library members that could grow on plates with ampicillin (with or without maltose)

Ampicillin (μg/ml)	Maltose? (5 mM)	T164-165 DKS	T164-165 GSGGG	EE DKS	EE GSGGG	Random DKS	Random GSGGG
5	no	734	878	7052	3510	nd	2458
50	no	394	294	1747	1159	nd	783
200	no	220	nd	1080	298	nd	nd
1000	no	nd	74	nd	nd	nd	60
5	yes	1098	761	8354	4056	nd	1969
50	yes	515	361	2414	1615	191	1156
200	yes	182	240	1525	630	nd	272
1000	yes	nd	88	nd	nd	nd	34

EE = end-to-end (insertion at C-terminus); nd = not determined

Table 3 Number of library members screened (picked from plates with indicated ampicillin and maltose levels)

Ampicillin (μg/ml)	Maltose? (5 mM)	T164-165 DKS	T164-165 GSGGG	EE DKS	EE GSGGG	Random DKS	Random GSGGG
5	no	-	96	-	288	-	96
50	no	-	-	-	-	-	-
200	no	-	-	-	-	-	480
1000	no	-	-	-	-	-	-
5	yes	96	192	-	864	-	768
50	yes	672	576	576	768	384	960
200	yes	80	384	-	-	-	1008
1000	yes	-	-	-	-	-	-

EE = end-to-end (insertion at C-terminus)

Table 4. Selected BLA-MBP Molecular Switches

Switch	Sequence	Switching effect*
IFG-5-277	MBP[1-165]-BLA[218-286]-GSGGG-BLA[24-215]-MBP[164-370]	-250%
IFD-5-7	MBP[1-165]-BLA[110-286]-DKS-BLA[24-107]-MBP[164-370]	+96%
IFD-5-15	MBP[1-165]-BLA[168-286]-DKS-BLA[24-170]-MBP[164-370]	+97%
EEG-50-530	MBP[1-370]-BLA[114-286]-GSGGG-BLA[24-112]-GSQQH	+228%
EEG-50-251	MBP[1-370]-BLA[114-286]-GSGGG-BLA[24-114]-K	+234%
RG-200-151	MBP[1-341]-A-BLA[111-286]-GSGGG-BLA[24-115]-MBP[347-370]	+350%
RG-200-604	MBP[1-331]-BLA[113-286]-GSGGG-BLA[24-112]-L-MBP[333-370]	+390%
RG-5-169	MBP[1-338]-BLA[34-286]-GSGGG-BLA[24-29]-MBP[337-370]	+855%
RG-200-13	MBP[1-316]-BLA[227-286]-GSGGG-BLA[24-226]-S-MBP[319-370]	+1650%

* Percent change in velocity of nitrocefin hydrolysis (50 μM nitrocefin) in the presence of 5 mM maltose in 100 mM sodium phosphate buffer, pH 7.0.

Table 5. Sugar dependence of switching effect of RG-200-13*.

Sugar	Binds to MBP?	Change in velocity of nitrocefin hydrolysis in presence of sugar
Sucrose	No	-5%
Lactose	No	-4%
Galactose	No	-3%
Maltose	Yes	+1800%
Maltotriose	Yes	+1700%
Maltotetraitol	Yes	+400%
β-cyclodextrin	Yes	+150%

*50 μM nitrocefin, 100 mM sodium phosphate buffer, pH 7.0, 22°C, 5 mM sugar except for β-cyclodextrin (3mM).

Table 6 Kinetic parameters of nitrocefin hydrolysis of RG-200-13 molecular switch.

Substrate	k_{cat} (s^{-1})			K_m (μM)			k_{cat}/K_m
	No maltose	5 mM maltose	Ratio[a]	No maltose	5 mM maltose	Ratio[a]	Ratio[a]
nitrocefin	~80	~520	~6.5	~325	~85	~0.26	~25

[a](with maltose)/(without maltose). Conditions: 100 mM sodium phosphate buffer, pH 7.0, 22°C.

Table 7. Effect of maltose on other substrates of switch RG-200-13

Substrate	Substrate concentration	K_m for TEM-1 β-lactamase[a]	Approximate fold increase in velocity of nitrocefin hydrolysis in the presence of 5 mM maltose
cephalothin	250 μM	246 μM	32
ampicillin	100 μM	32 μM	26
	500 μM		10
benzylpenicillin	100 μM	19 μM	17
	500 μM		7
carbenicillin	1 mM	?	4
oxacillin	1 mM	3 μM	5

Conditions:100 mM sodium phosphate buffer, pH 7.0, 22°C. [a](Raquet, Lamotte-Brasseur et al. 1994)

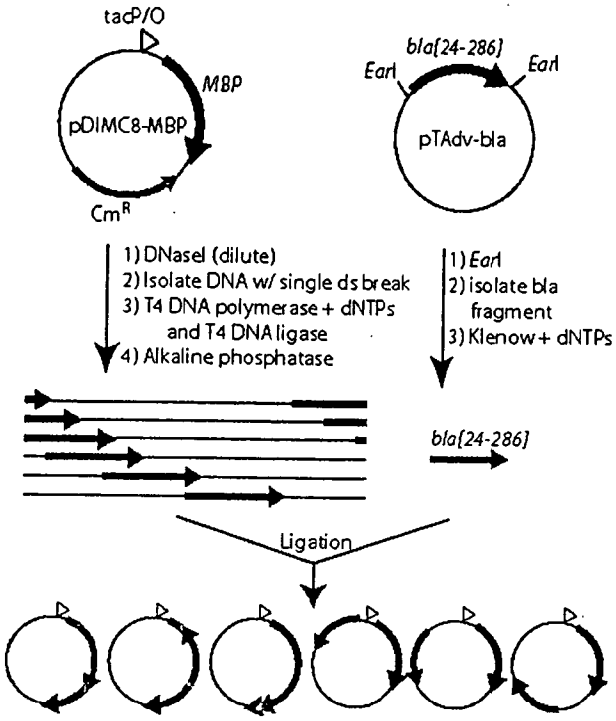

Figure 1 Random Domain Insertion.

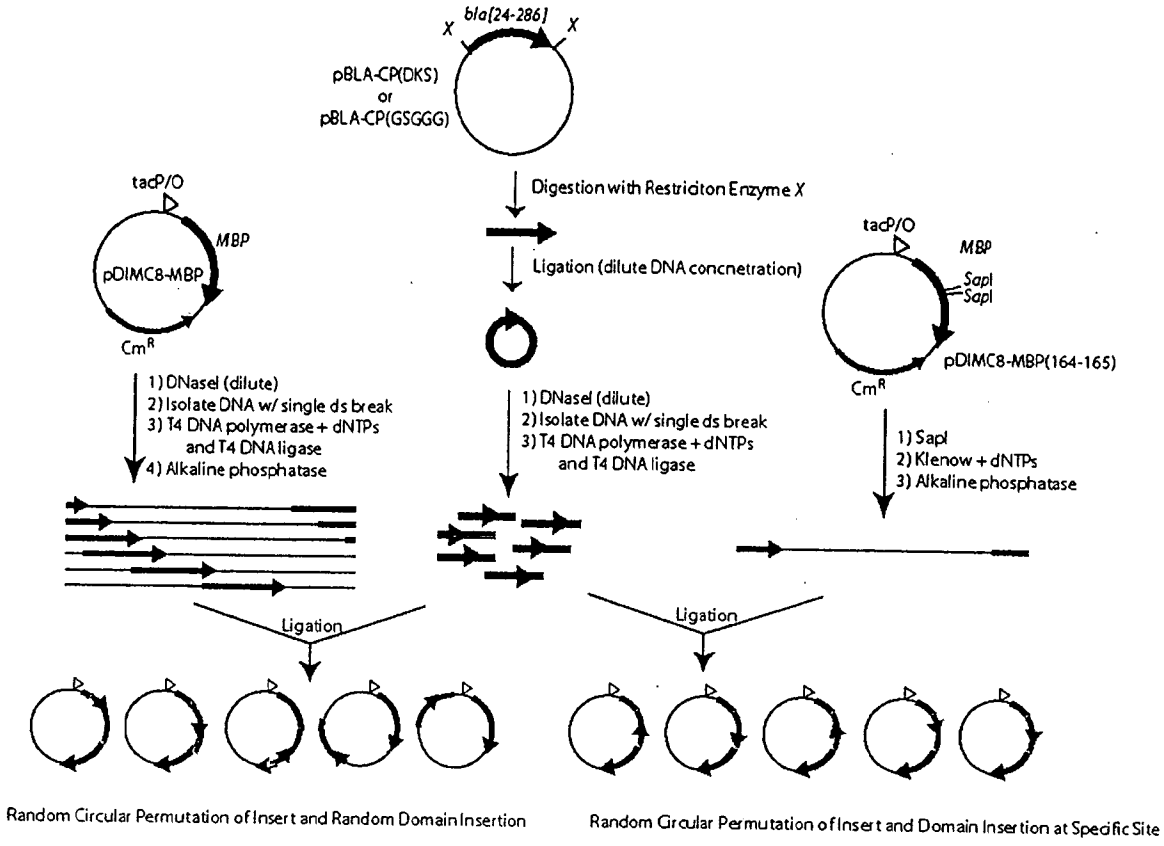

Figure 2 Creation of domain insertion libraries involving the random circular permutation of DNA.

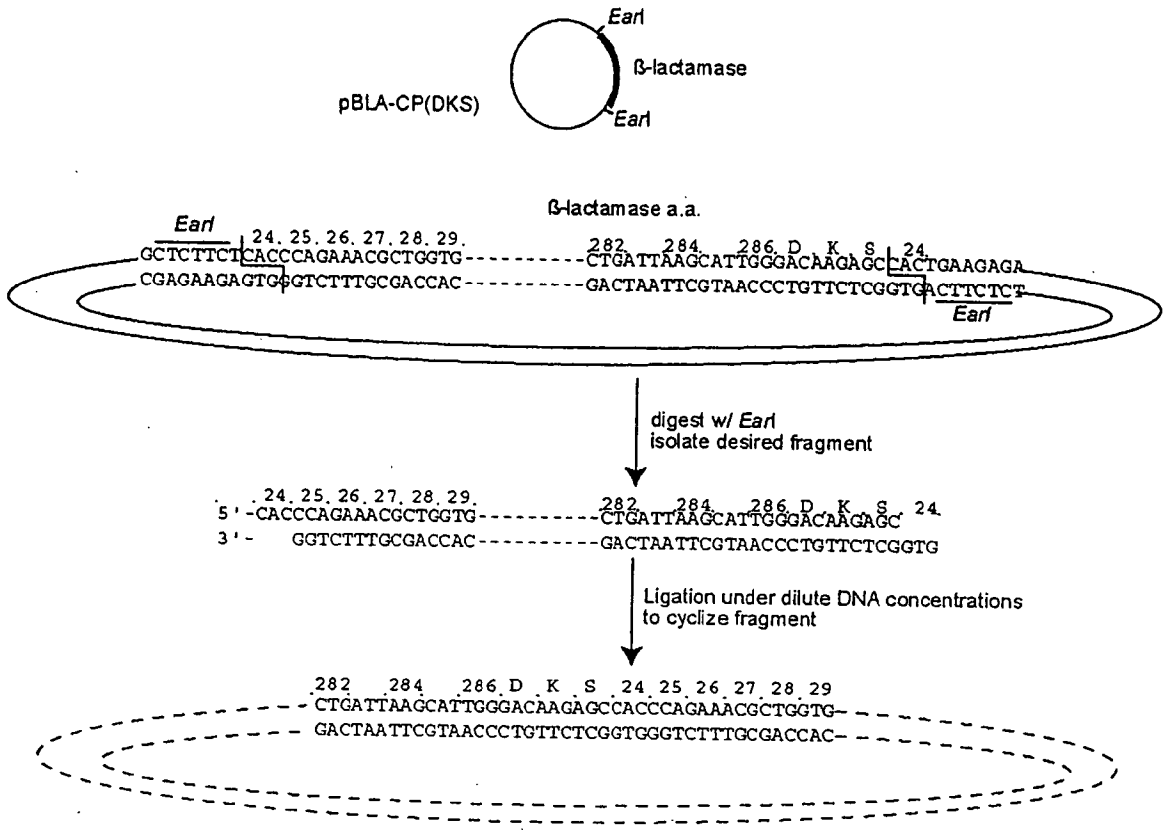

Figure 3 Creation of a cyclized beta-lactamase gene with a DKS linker.

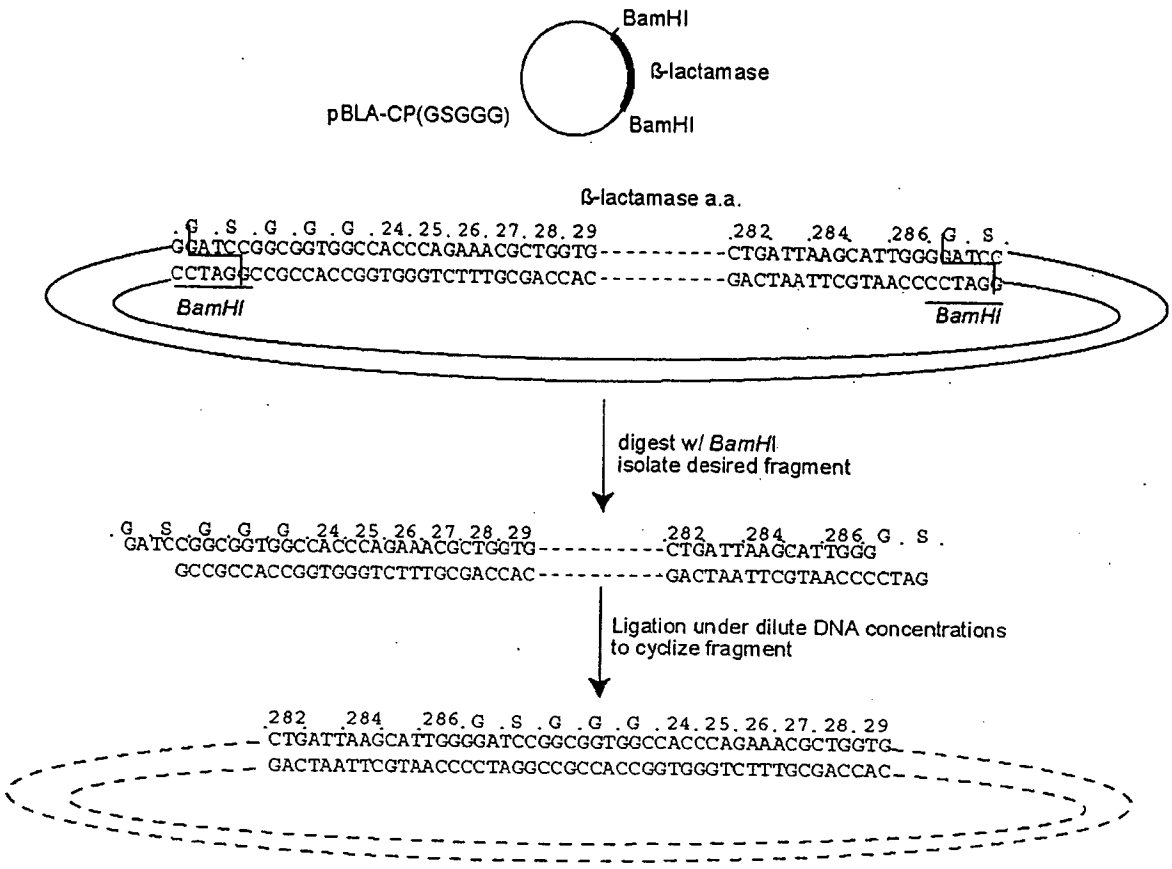

Figure 4 Creation of a cyclized beta-lactamase gene with a GSGGG linker.

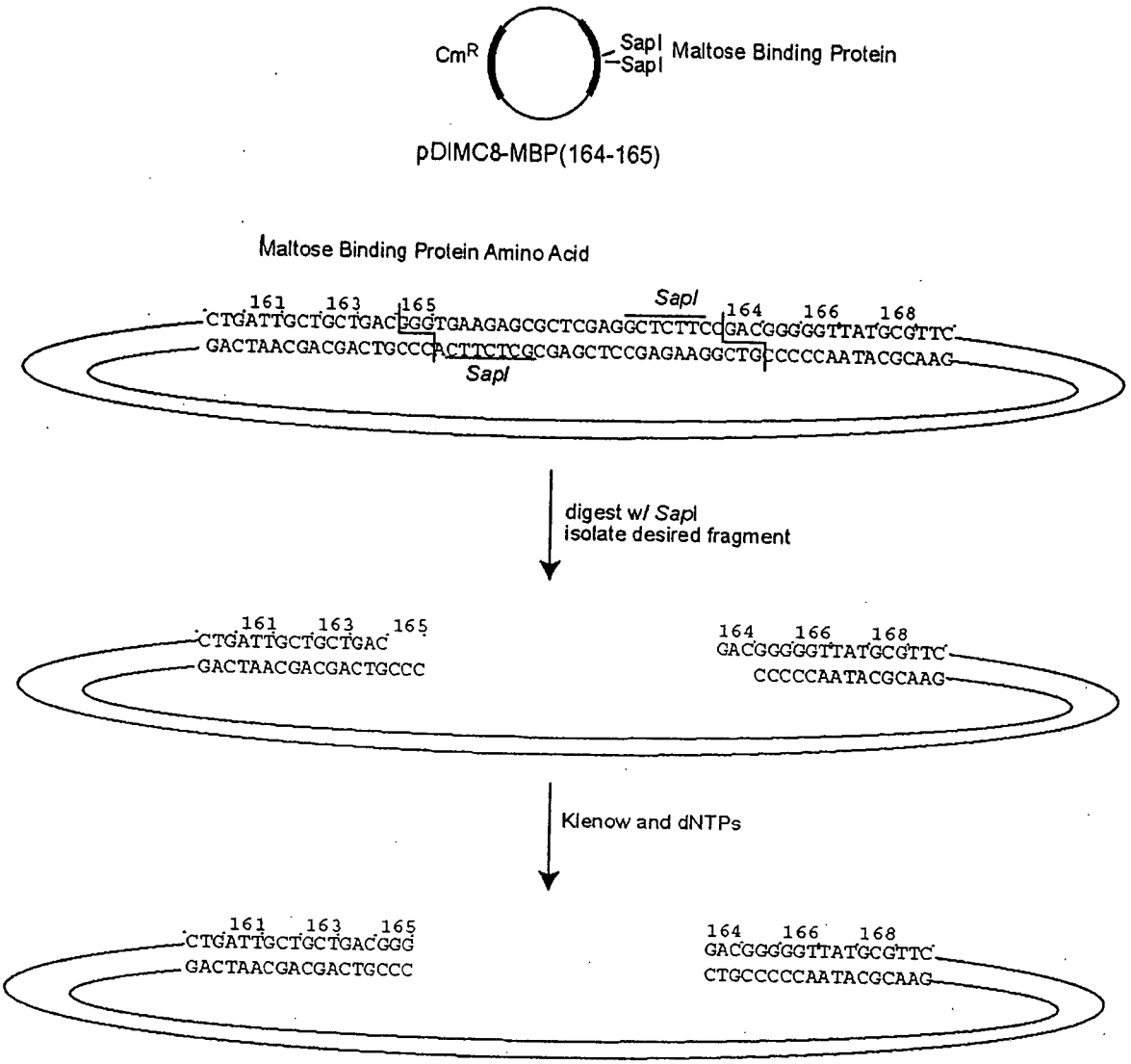

Figure 5 Prepartion of target DNA for insertion of DNA at a specific location (in this case between MBP [1-165] and MBP [164-370])

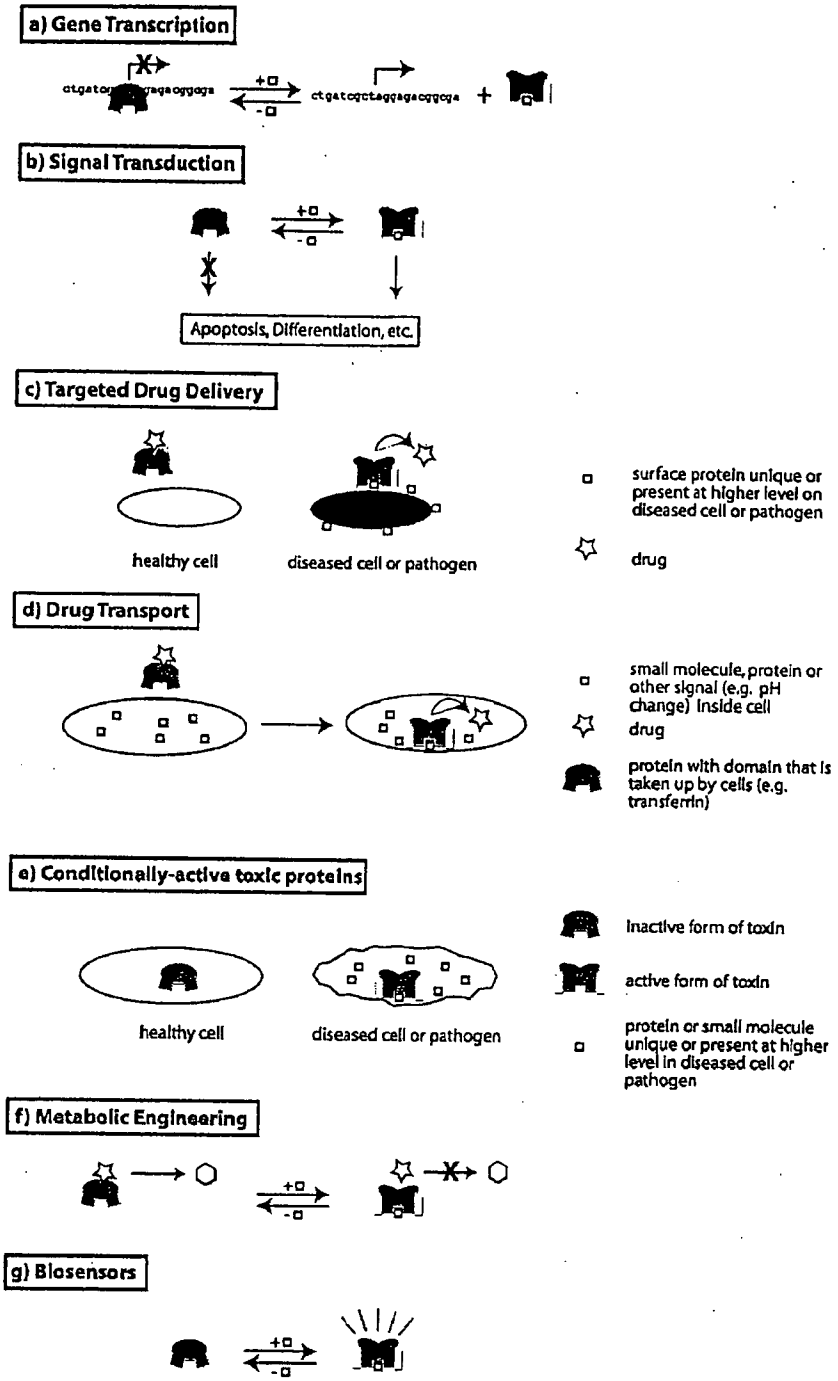

Figure 6 Application of Molecular Switches. (a) Regulation of gene transcription. A DNA binding protein is functionally coupled to a protein that binds ❐ such that the protein releases from the DNA in the presence of ❐ . The signaling molecule ❐ , which could be a naturally occurring in the cell or supplied as a drug, could then be used to induce expression of an exogenously supplied therapeutic or toxic protein or to repress/induce expression of endogenous genes in order to restore normal cell function. (b) Modulation of cell signaling pathways. Cell fates can be regulated by engineering a protein involved in a cell signaling pathways to be conditionally active in the presence of the signal ❐ , which could be a naturally occurring small molecule, a

protein or an administered drug. (c) Targeted drug delivery. A drug binding protein can be functionally coupled to a protein that binds some small molecule or protein that is a signature of a disease (e.g. a surface protein). The drug binding protein releases the drug only in the presence of the molecular signature of the disease. (d) Drug transport. A drug binding protein can be fused to a protein that is taken up by cells in order to transport the drug into cells. The two proteins are functionally coupled such that the conformational change in the transport protein upon interaction with ❐ inside the cell is coupled to the release of the drug. (e) *Conditionally active toxic proteins*. A toxic protein is engineered such that it is nontoxic unless it interacts with ❐ , which is a small molecule or protein signature of a diseased cell or pathogen. (f) *Metabolic engineering*. Just as small molecules in cells can allosterically modulate metabolic pathways, engineered molecular switches can control metabolic pathways. This could be achieved by functional coupling of an enzyme (whose activity is to be modulated) with a protein that binds ❐ . (g) *Biosensors*. The key design issue in biosensors is the creation of a macromolecule that specifically binds its target ligand and, upon binding, transduces this into a signal (such as optical or electrochemical changes) that can be detected macroscopically. The functional coupling of a binding protein and a signal transducing properties is a general strategy towards engineering new families of biosensors.

Abstract submitted for the ACS National Meeting

The combinatorial design of protein molecular switches

Marc Ostermeier, and Gurkan Guntas, Department of Chemical and Biomolecular Engineering, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, Fax: 410-516-5510, oster@jhu.edu

A hallmark of biological systems is the high degree of interactions amongst and within their constituent components. One advantage that such interactions bring is the establishment of coupling between different functions. A protein that couples two functions can be described as a molecular switch. For example, an allosteric enzyme is a switch that couples effector levels (input) to enzyme activity (output). In most general terms, a molecular switch couples signals (e.g. ligand binding, protein-protein interactions, pH, covalent modification, temperature) to functionality (e.g. enzymatic activity, binding affinity, fluorescence). Molecular switches can be of an "on/off" nature or such that the signal modulates the function between two different levels of activity. The network of such molecular switches establishes the complex circuits that control cellular processes. Protein molecular switches have a wide variety of potential applications including the regulation of gene transcription, the modulation of cell signaling pathways, targeted drug delivery, drug transport, the creation of conditionally active toxic proteins, metabolic engineering, and the creation of novel biosensors. We have devised a combinatorial protein engineering algorithm that can create molecular switches by the coupling of two proteins' functionalities. We have demonstrated the efficacy of the algorithm by creating a set of allosteric enzymes from two unrelated proteins with the prerequisite effector-binding and catalytic functionalities, respectively. In some of these engineered molecular switches, the presence of effector increased catalytic activity by more than ten-fold.

37

Described are methods for creating and using molecular switches with different affinities for the effector molecule (signal) as well as methods for creating and using molecular switches with different specificities for the effector molecule (i.e. it responds to a new signal).

We have demonstrated a method for creating a molecular switch by a combinatorial approach of recombining two genes so as to couple their functions. We have demonstrated the approach by recombining the genes for TEM1 β-lactamase (BLA) and the *E. coli* maltose binding protein (MBP) such that BLA catalytic activity was a function of the concentration of maltose. The kinetic parameters and affinity for maltose of one of these switches (RG13 –referred to in the previous provisional as RG-200-13 – which has the amino acid sequence MBP[1-316]-BLA[227-286-GSGGG-BLA[24-226-S-MBP[319-370]) have been determined and are shown in Tables 1 and 2. RG13 has a dissociation constant for maltose is about 5-6 μM in the absence of a BLA substrate. In the presence of saturating amounts of the substrate carbenicillin, the dissociation constant for maltose decreases to about 1 μM. This shows that the binding of maltose and substrate (carbenicillin) are coupled. The coupling energy is on the order of 1 kcal/mol. This is consistent with a decrease in K_m for nitrocefin in the presence of maltose (Table 2)

A switch is most useful if the range of the concentration of the signal (maltose, in the case of RG13) overlaps with the range of concentration of the signal that the dependent function responds to. For the case when a ligand-binding protein is used as the signal detector and the ligand is the signal, the later range corresponds to approximately to the range $0.1K_d - 10\,K_d$, where K_d is the dissociation constant of the switch and the signal. This can be seen from the following example:

In the case of RG13, the velocity of nitrocefin hydrolysis is the dependent function. The velocity (v) of nitrocefin hydrolysis depends on the steady state kinetic parameters by the Michaelis-Menten (Equation 1).

$$v = \frac{[E]_0[S]k_{cat}}{K_m + [S]} \quad (1)$$

where $[E]_0$ is the concentration of the switch, [S] is the concentration of nitrocefin and k_{cat} and K_m are the Michaelis-Menten kinetic parameters. In the absence of maltose, the velocity is found by Equation 2

$$v^- = \frac{[E]_0[S]k^-_{cat}}{K^-_m + [S]} \quad (2)$$

where the superscript "-" designates that the parameters are those when maltose is not bound to the switch. In the presence of saturating concentrations of maltose (i.e. maltose is bound to all switches; this occurs at very high concentrations of maltose relative to the dissociation constant K_d for maltose), the velocity is found by Equation 3

$$v^+ = \frac{[E]_0[S]k^+_{cat}}{K^+_m + [S]} \quad (3)$$

where the superscript "+" designates that the parameters are those when maltose is bound to the switch. At intermediate concentrations of maltose, the velocity depends on the fraction of switches that have maltose bound. If we make the approximation that the small cooperative effect of

maltose- and substrate-binding can be ignored, the fraction F of switches that are bound to maltose can be found by Equation 4.

$$F = \frac{[M]}{[K_d]+[M]} \quad (4)$$

where [M] is the concentration of maltose. The velocity of nitrocefin hydrolysis is thus found by Equation 5

$$v = F\frac{[E]_0[S]k_{cat}^{+}}{K_m^{+}+[S]} + (1-F)\frac{[E]_0[S]k_{cat}^{-}}{K_m^{-}+[S]} \quad (5)$$

Equation 5 is true for all concentrations of maltose as it reduces to Equations 2 and 3 in the limiting cases of no maltose bound and saturating maltose, respectively. The fold-increase in the rate of nitrocefin velocity Z is found by dividing the right hand side of Equation 5 by the velocity in the absence of maltose to get Equation 6

$$Z = F\frac{k_{cat}^{+}\left(K_m^{-}+[S]\right)}{k_{cat}^{-}\left(K_m^{+}+[S]\right)} + (1-F) \quad (6)$$

Equation 6 is plotted in Figure 1 for the case of RG13 hydrolysis of 25 μM nitrocefin using a range of different dissociation constants for maltose. One can see that the velocity is changing most in the range of one order of magnitude higher or lower than the dissociation constant for maltose. The switch will have the largest change in the dependent function if the concentration of the signal (in the case of RG13 this is maltose) changes within this range or changes through this range. Thus, it is desirable for the application of molecular switches to create switches with different affinities for the signal so as to be useful for different concentration ranges of the signal.

In one aspect of the invention, switches created by the method have different affinities for maltose. For example switch RG-5-169 (sequence MBP[1-338]-BLA[34-286]-GSGGG-BLA[24-29]-MBP[337-370] has a K_d for maltose (> 1 mM) that is much greater than RG13 has for maltose (1-5 μM).

In another aspect of the invention, the effector-affinity of switches is altered by a variety of methods including rational design and directed evolution methods. As long as the resulting altered-affinity switch maintains a conformational change upon binding the effector that results in changes the dependent function, switching will be maintained.

In another aspect of the invention mutations already known to alter the affinity of the ligand recognition domain (for RG13 this is MBP) are introduced into the switch to create switches with altered affinity for the ligand. These mutations consist of those that make direct contact with the ligand, those that make contact with residues that make direct contact with the ligand and those that are more distal from the binding site pocket. For example, Marvin and Hellinga (Marvin and Hellinga 2001) have described a set of mutations in the hinge region of MBP that manipulate the conformational equilibria between the open and closed state (Marvin and Hellinga 2001). Residual dipolar couplings have been used to establish that the apo forms of these mutants are partially closed relative to the apo wildtype MBP with the closure angles being 9.5° and 28.4° for I329W and A96W/I329W, respectively (the ligand-bound closed form of MBP has a closure angle of 35°)

(Millet, Hudson et al. 2003). Because partial closing shifts the equilibrium towards the ligand-bound state, the I329W mutation results in about a 20-fold increase in affinity for maltose and the A96W/I329W double mutant results in a 60-fold increase in the affinity for maltose compared to wildtype MBP at 25°C (Marvin and Hellinga 2001). The affinities of MBP, MBP(I329W) and MBP(I329W/A96W) are 800 nM, 35 nM and 13 nM, respectively. Introduction of these mutations into RG13 results in mutants with increased affinity for maltose (Table 1) while still maintaining switching behavior (Table 2 and Figure 2). In addition, the level of activity in the presence of saturating amounts of maltose (the "on" state) was not affected by the mutations (Table 1). From a practical standpoint, the increase in maltose affinity of these hinge mutants indicates that ligand-affinity of RG13 can be systematically changed to create molecular switches that respond to different concentration ranges of effector while still maintaining switching ability and high activity in the presence of the effector. By increasing the affinity for maltose one increases the sensitivity of the switch (i.e. it will switch to a higher level of activity at lower concentrations of maltose). In a further aspect of the invention, combinations of these affinity-altered switches behave as a composite switch with a high dynamic range.

The invention also consists of methods to alter the specificity of the signal recognition domain so that it recognizes other signals. This would allow the construction of new molecular switches in which the dependent function responded to new signals without requiring the construction of entirely new molecular switches. For the example of RG13, where the signal binding domain is the maltose binding protein, these methods can change the ligand to which the switch binds to. This would allow the construction of molecular switches in which BLA activity could be modulated by different ligands without developing entirely new molecular switches. Thus, in a further aspect of the invention, the identity of the signal to which the switch responds can be altered by mutations in existing switches. For example, mutations in the signal recognition domain already known to alter the ligand-binding specificity can be introduced into the switch to create switches that respond to new ligands. For example, Hellinga and colleagues have computationally designed periplasmic binding proteins with radically different binding specificities (Looger, Dwyer et al. 2003) including designing MBP to bind Zn^{2+} (Marvin and Hellinga 2001) instead of maltose. MBP binds maltose with high affinity (K_d = 0.8 μM) but does not bind Zn^{2+}. MBP with the A* set of mutations (A63H/R66H/Y155E/W340E) has high affinity for Zn^{2+} (K_d = 5.1 μM) and does not bind maltose (Marvin and Hellinga 2001). Introduction of the A* set of mutations into RG13 results in a switch that responds to Zn^{2+} but not maltose.

In a further aspect of the invention, the signal recognition domain is altered by rational design or directed evolution to bind to new effectors. With respect to testing mutations predicted by rationale design or screening or selecting libraries created for a directed evolution approach, existing switches are used to efficiently test or select for binding to new ligands *in vivo*. For example, *E. coli* cells expressing the MBP-BLA switch RG13 from the lac promoter on pDIMC8 have a higher MIC for ampicillin (Amp) in the presence of maltose than in their absence (Table 3) because the BLA enzymatic activity of RG13 (hydrolysis of ampicillin) is higher in the presence of maltose. Thus, for example, mutations created in RG13 (either by rational design or by a stochastic or semi-stochastic method) such that mutant forms of RG13 bind another ligand X (and behave as a switch) can be screened or selected for *in vivo*. *E. coli* producing such a new switch will grow at 200 μg/ml Amp in the presence of X but not in the absence of X.

In a further aspect of the invention, ligands that bind to the signal recognition domain in a different manner have different switching ability. This is demonstrated by the fact that

β-cyclodextrin, which is known to bind to MBP but with a different conformational change (Skrynnikov, Goto et al. 2000; Hwang, Skrynnikov et al. 2001), changes the activity of the RG13 switch in a different manner than maltose (see Table 2).

Table 1 Hinge Mutations in MBP domain Increase Maltose Affinity in RG13.

Protein	Ligand	K_d maltose (μM)[a]			
		No Substrate[b]	25 μM nitrocefin[c]	Saturating Carbenicillin	
				IPF[b]	Enzymatic assay[d]
RG13	maltose	5.5 ± 0.5	6.7 ± 0.03	1.3 ± 0.5	0.9 ± 0.1
RG13 I329W	maltose	0.55 ± 0.13	1.0 ± 0.04	nd	0.11 ± 0.01
RG13 I329W/A96W	maltose	nd	0.17 ± 0.02	nd	nd

[a]Conditions: 100 mM NaCl, 50 mM $NaPO_4$, pH 7.0, 25°C

[b]Determined by measuring intrinsic protein fluorescence (IPF) as a function of maltose concentration. When using IPF at saturating carbenicillin, a concentration of 10 mM carbenicillin was used.

[c]Determined by measuring the initial rate of nitrocefin hydrolysis as a function of maltose concentration. 25 μM nitrocefin is well below the K_m of nitrocefin. Thus, most molecules of RG13 will not have nitrocefin bound and the effective K_d that is measured is close to what it would be in the absence of substrate.

[d]Determined by measuring the initial rate of carbenicillin hydrolysis as a function of maltose concentration. A concentration of 1.5 mM carbenicillin was used, which is well above the K_m of carbenicillin. Thus, most molecules of RG13 will have carbenicillin bound and the K_d that is measured is in the presence of bound substrate (carbenicillin).

Table 2 Kinetic parameters of nitrocefin hydrolysis[a].

Protein	Effector	k_{cat} (s^{-1})[a]	k_{cat} Ratio[b]	K_m (μM)[a]	K_m Ratio[b]	k_{cat}/K_m[a] (s^{-1} μM^{-1})	k_{cat}/K_m Ratio[b]
RG13	-	200 ± 40	-	550 ± 120		0.37 ± 0.10	
RG13 I329W	-	190 ± 30	-	350 ± 60		0.54 ± 0.11	
RG13 I329W/A96W	-	360 ± 40	-	260 ± 40		1.4 ± 0.3	
RG13	maltose	620 ± 30	3.1 ± 0.6	68 ± 4	0.12 ± 0.03	9.2 ± 0.7	25 ± 7
RG13 I329W	maltose	590 ± 50	3.1 ± 0.5	53 ± 7	0.15 ± 0.03	11.0 ± 1.8	20 ± 5
RG13 I329W/A96W	maltose	530 ± 20	1.5 ± 0.2	60 ± 4	0.23 ± 0.04	8.9 ± 0.8	6.4 ± 1.3
RG13	β-cyclo[c]	590 ± 60	2.9 ± 0.6	870 ± 90	1.6 ± 0.4	0.67 ± 0.10	1.8 ± 0.6

[a]Conditions: 100 mM sodium phosphate buffer, pH 7.0, 25°C; concentration of effector is 5 mM

[b](with effector)/(without effector).

[c]β-cyclodextrin

Table 3 Minimum inhibitory concentration of ampicillin for *E. coli* cells expressing RG13[a]

Supplement to plate	MIC ampicillin (μg/ml)
none	100
50 μM maltose	400
5 mM maltose	400

[a]conditions: LB plates, 37 °C, supplemented with maltose as indicated. Approximately 1000 colony forming units (without ampicillin) per plate. Concentrations of ampicillin tested: 0, 25, 50, 100, 200, 400 and 800 μg/ml.

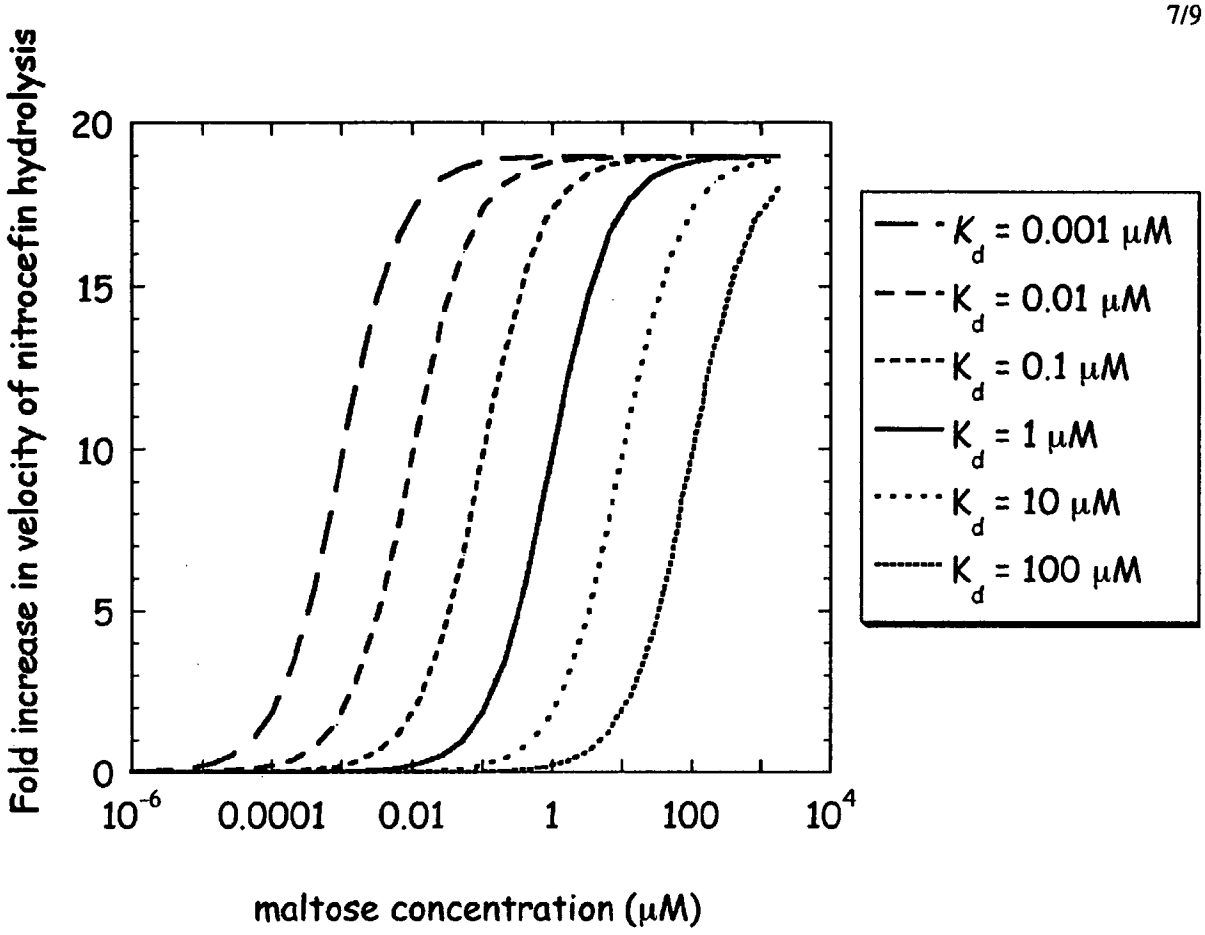

Figure 1 Fold increase in velocity of switch RG13 with different dissociation concentrations (K_d) for maltose. Concentration of nitrocefin was 25 μM. The kinetic parameters of RG13 with and without maltose are those in Table 2. Equation 6 was used to generate the curves.

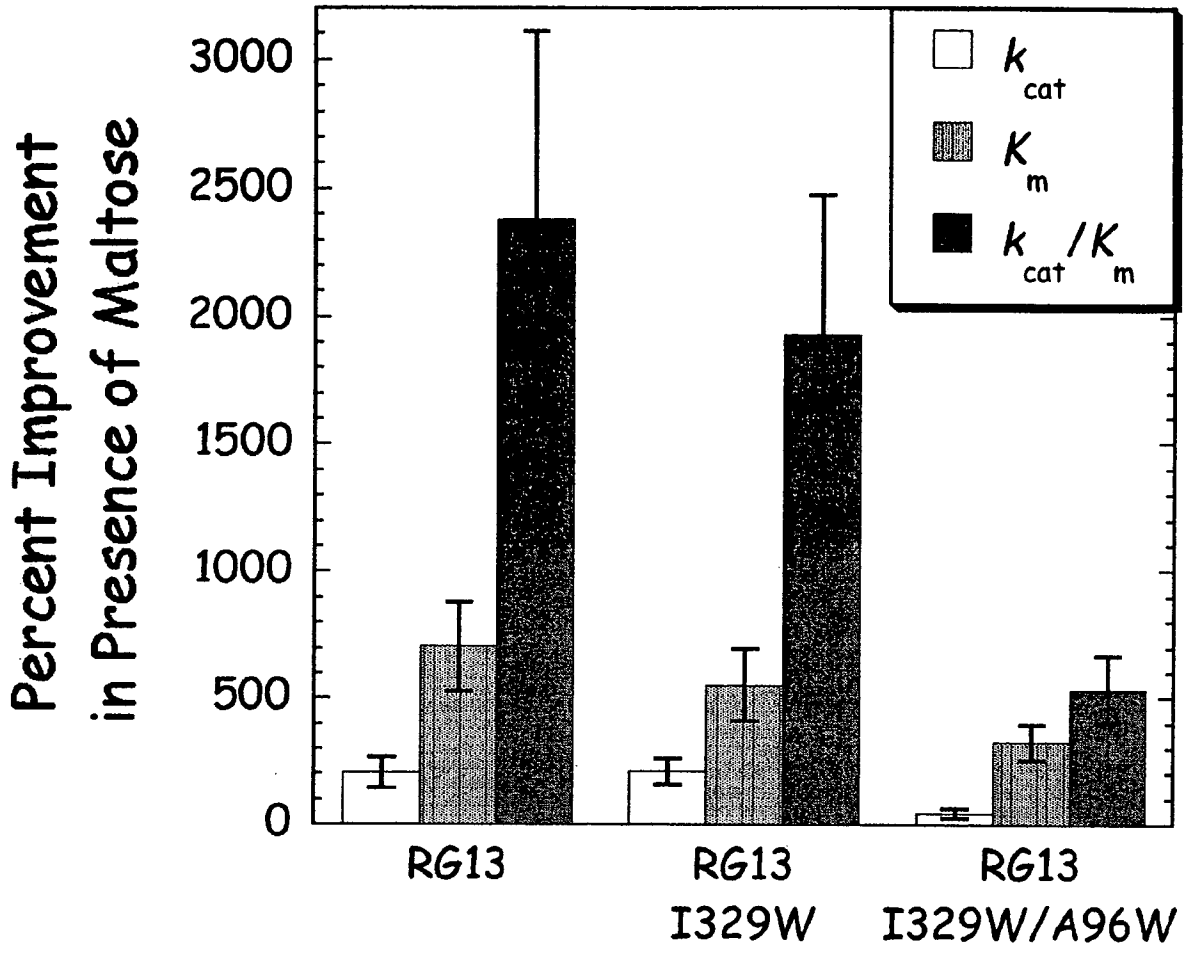

Figure 2. Percent improvement in kinetic parameters of nitrocefin hydrolysis of switches in presence of maltose. Data taken from Table 1. For k_{cat} and k_{cat}/K_m, an improvement is an increase in the value. For K_m, an improvement is a decrease in the value.

References

Hwang, P. M., N. R. Skrynnikov, et al. (2001). "Domain orientation in beta-cyclodextrin-loaded maltose binding protein: diffusion anisotropy measurements confirm the results of a dipolar coupling study." J Biomol NMR **20**(1): 83-8.

Looger, L. L., M. A. Dwyer, et al. (2003). "Computational design of receptor and sensor proteins with novel functions." Nature **423**(6936): 185-90.

Marvin, J. S. and H. W. Hellinga (2001). "Conversion of a maltose receptor into a zinc biosensor by computational design." Proc Natl Acad Sci U S A **98**(9): 4955-60.

Marvin, J. S. and H. W. Hellinga (2001). "Manipulation of ligand binding affinity by exploitation of conformational coupling." Nat Struct Biol **8**(9): 795-8.

Millet, O., R. P. Hudson, et al. (2003). "The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy." Proc Natl Acad Sci U S A **100**(22): 12700-5.

Skrynnikov, N. R., N. K. Goto, et al. (2000). "Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: differences in solution and crystal forms of maltodextrin binding protein loaded with beta-cyclodextrin." J Mol Biol **295**(5): 1265-73.

A Molecular Switch Created by *In Vitro* Recombination of Non-Homologous Genes

Gurkan Guntas[1], Sarah F. Mitchell[2], and Marc Ostermeier[1,*]

[1]Department of Chemical and Biomolecular Engineering
[2]Program in Molecular Biophysics
Johns Hopkins University
3400 N. Charles St.
Baltimore, Maryland 21218

Tel: 410.516.7144
Fax: 410.516.5510
*Correspondence: oster@jhu.edu

Summary

We have created a molecular switch by the *in vitro* recombination of non-homologous genes and subjecting the recombined genes to evolutionary pressure. The gene encoding TEM1 β-lactamase was circularly permuted in a random fashion and subsequently randomly inserted into the gene encoding *Escherichia coli* maltose binding protein. From this library a switch (RG13) was identified in which its β-lactam hydrolysis activity was compromised in the absence of maltose but increased 25-fold in the presence of maltose. Upon removal of maltose, RG13's catalytic activity returned to its pre-maltose level, illustrating that the switching is reversible. The modularity of RG13 was demonstrated by increasing maltose affinity while preserving switching activity. RG13 gave rise to a novel cellular phenotype, illustrating the potential of molecular switches to rewire the cellular circuitry.

Running Title: Creation of a molecular switch by recombination

Introduction

The incredible complexity of biological systems derives to a large extent from the high degree of interactions amongst their constituent components. As such, the cell is often described as a complex circuit consisting of an interacting network of molecules. A key component of these networks are protein "switches" that serve to couple cellular functions. A switch recognizes an input signal (e.g. ligand concentration, pH, covalent modification) and, as a result, its output signal (e.g. enzyme activity, ligand affinity, oligomeric state) is modified. Examples of natural switches include allosteric enzymes which couple effector levels to enzymatic activity and ligand-dependent transcription factors that couple ligand concentration to gene expression. The ability to create novel switches or to modify existing switches by coupling hitherto uncoupled functions would enable the rewiring of the cellular circuitry to our own design. In addition, the ability to create protein switches has tremendous practical potential for developing novel molecular sensors, "smart" materials, and as a tool for elucidating molecular and cellular function.

We sought a general approach for creating switches in which one could select from natural or engineered proteins with the desired input and output functions and, by combining the proteins in a systematic fashion, create switches in which their functions were tightly coupled. In addition, we sought switches in which the presence of the signal resulted in an increase in the output function. Existing approaches [1-16] either require existing switches as starting points, are limited to specific input functions and mechanisms, lack reversibility, or lack practicality *in vitro* or *in vivo*. In addition, most

existing approaches have not been shown to produce switches with large changes in output (>10-fold) in response to a signal.

Our approach involves recombination between two genes encoding the prerequisite input and output functions. Such an approach is inspired by the evolutionary mechanism of domain recombination [17] – a major facilitator in the natural evolution of protein function [18]. Domain recombination strategies have been applied to the construction of switches [2, 7, 10-16], but the very limited manner in which the domains were recombined has restricted the success of this approach. We reasoned that a more diverse exploration of fusion geometries between two proteins would enable the creation of switches with superior properties. The structural space that we sought to explore can be conceptualized as "rolling" the two proteins across each others surface and fusing them through peptide bonds at the points where their surfaces meet. We developed a homology-independent, combinatorial method for recombining genes that samples such a structural space. The method involves two basic steps. First, the gene to be inserted is circularly permuted in a random fashion to vary the fusion site on the insert gene. Second, the library of the circularly permuted gene is randomly inserted into the target gene.

Results and Discussion

We applied this approach to the recombination of the genes encoding TEM-1 β-lactamase (BLA) and the *E. coli* maltose binding protein (MBP). BLA and MBP lack any sequence, structural or functional relationship except for the fact that they are

periplasmic proteins of bacterial origin. BLA is a monomeric enzyme that hydrolyzes the amide bond of the β-lactam ring of β-lactam antibiotics. The presence of maltose has no effect on wild type BLA enzymatic activity, with or without the presence of an equimolar amount of MBP [15]. MBP is a member of the periplasmic binding protein superfamily and is involved in chemotactic response and the transport of maltodextrins. MBP consists of a single polypeptide chain that folds into two domains connected by a hinge region. The single binding site for maltose is at the interface of these two domains. In the absence of maltose, MBP exists in an open form. Maltose-binding is concomitant with a 35° bending motion about the hinge resulting in the closed form of the protein [19]. We sought to create a molecular switch by combining BLA and MBP in such a manner that the rate of β-lactam hydrolysis was coupled to maltose-binding and maltose concentration. We reasoned that in such a switch the conformational change in the MBP domain upon maltose binding would propagate to the active site of the BLA domain and alter its catalytic properties, a mechanism analogous to natural allosteric effects.

The fragment of the *BLA* gene coding for the mature protein was circularly permuted in a random fashion [20, 21] and subsequently randomly inserted [15] into a plasmid containing the *E. coli malE* gene that codes for MBP (Figure 1A). For the random circular permutation of *bla[24-286]*, we fused the 5' and 3' ends by an oligonucleotide sequence that would result in a GSGGG flexible peptide linker between the original N- and C- termini of the protein. This linker was designed to be of sufficient length to connect the termini without perturbing BLA structure. Statistical analysis on the resulting library indicated that a minimum of 27,000 members contained a circularly

permuted *bla[24-286]* inserted into *malE* in the correct orientation with both fusion points in-frame with *malE*. Approximately 0.33% of these members were able to form colonies on rich media plates containing 200 μg/ml ampicillin and 50 mM maltose. These library members were screened in 96-well format for a maltose dependence on β-lactamase activity using a colorimetric assay for nitrocefin hydrolysis. We identified one protein (RG13; Figure 1B) in which the initial velocity of nitrocefin hydrolysis (at 50 μM nitrocefin) increased by 17-fold in the presence of maltose. In RG13, the BLA was circularly permuted in a loop that precedes a β-sheet that lines the active site of the enzyme. The circular permuted BLA was inserted at the beginning of an α-helix of MBP such that two MBP residues were deleted (Figure 1C).

Using purified RG13, we confirmed that the increase in catalytic activity occurred only in the presence of sugars that are known to bind and induce a conformational change in MBP (Figure 2A). Sugars known to induce a large conformational change [22] (maltose and maltotriose; 35° closure angle) produced a 15- to 20-fold increase in the rate of nitrocefin hydrolysis. β-cyclodextrin, which only induces a 10° hinge bending motion in MBP [24], increased the rate 2-fold. Non-ligands such as sucrose, lactose and galactose had no effect. We next determined that the switching was reversible (i.e. upon removing maltose, the activity returns to its pre-maltose level). This was demonstrated both by competing maltose off RG13 using β-cyclodextrin (Figure 2B) and by subjecting RG13 to repeated rounds of dialysis and addition of maltose to cycle between low and high levels of enzymatic activity (Figure 2C). This reversibility is one of the features that differentiates our approach from methods such as conditional protein splicing [8, 16] that

produce non-reversible switches that control the *production* of active protein rather than a protein's activity *per se*.

From steady state kinetics experiments, we determined RG13's Michaelis-Menten parameters for nitrocefin hydrolysis at 25°C in the absence and presence of maltose. In the absence of maltose, the catalytic constants were $k_{cat} = 200 \pm 40$ s^{-1} and $K_m = 550 \pm 120$ μM. With the addition of saturating amounts of maltose, k_{cat} increased 3-fold and K_m decreased 8-fold, resulting in a 25-fold increase in k_{cat}/K_m. The kinetic constants of RG13 in the presence of saturating concentrations of maltose ($k_{cat} = 620 \pm 60$ s^{-1} and $K_m = 68 \pm 4$ μM) were comparable to that previously reported for BLA at 24°C ($k_{cat} = 900$ s^{-1} and $K_m = 110$ μM; [25]) indicating that RG13 is a very active TEM1 β-lactamase in the presence of maltose. RG13 has exhibited switching behavior with all seven BLA substrate tested to date including ampicillin (16-fold rate increase at 50 μM ampicillin) and carbenicillin (12-fold rate increase at 50 μM carbenicillin).

The increase in k_{cat} indicates that maltose binding affects the catalytic steps. However, since K_m is a combination of the rate constants for substrate binding as well as catalysis [26], K_m could not be directly used to ascertain the effect of maltose on substrate binding. Instead, the effect of maltose on substrate binding was determined indirectly by measuring the effect of substrate on maltose binding using intrinsic protein fluorescence. These studies suggested that RG13 undergoes a conformational change much like MBP does upon maltodextrin-binding since maltose-induced quenching of total fluorescence (~10%) and shifting of the maximum fluorescence wavelength (a 1.5 nm red-shift for maltose and a 4 nm blue-shift for β-cyclodextrin) were similar to that previously reported

for MBP [27]. The presence of saturating amounts of the substrate carbenicillin decreased the dissociation constant of maltose and RG13 from 5.5 ± 0.5 μM to 1.3 ± 0.5 μM; thus, maltose binding must decrease the dissociation constant of carbenicillin and RG13 by the same factor (Supplementary Figure S1). This corresponds to a coupling energy of approximately 1 kcal/mol and offer an additional explanation for the increase in β-lactam hydrolysis in the presence of maltose: a positive heterotropic allosteric effect on substrate binding.

Presumably, the BLA domain of the apo, open form of RG13 exists in a compromised, less active conformation. In the ligand-bound state, the BLA domain exists in a more normal, active conformation. But what is the state of the BLA domain in the process of closing? At what closure angle do the catalytic properties of RG13 improve? To address these questions, we took advantage of mutations in the hinge region of MBP that manipulate the conformational equilibria between the open and closed state [28]. Residual dipolar couplings have been used to establish that the apo forms of these mutants are partially closed relative to the apo wildtype MBP with the ensemble average closure angles being 9.5° and 28.4° for I329W and I329W/A96W, respectively [29]. The ligand-bound closed forms of MBP, MBP(I329W) and MBP(I329W/A96W) have closure angles of 35°. Partial closing shifts the equilibrium towards the ligand-bound state and thus the mutations increase the affinity for maltose [28].

Introduction of these mutations into RG13 resulted in the creation of more sensitive switches–switches that respond to lower concentrations of maltose (Figures 3A

and 3B). The fact that we observed qualitatively similar changes in maltose-affinity when the mutations are introduced into RG13 strongly suggests that the relative order and magnitude of the angles of closure of RG13, RG13(I329W) and RG13(I329W/A96W) are similar to that of MBP, MBP(I329W) and MBP(I329W/A96W). Thus, the apo forms of the two RG13 mutants offer conformations intermediate between the open to the closed form of RG13–conformations that may reflect that of RG13 in the process of closing. Assuming that the process of closing in RG13 passes through the conformations of the apo forms of the two RG13 mutants, kinetic characterization of RG13(I329W) and RG13(I329W/A96W) suggested that the initial stages of closing do not result in changes in the BLA domain that substantially affect catalysis (Figures 3B-D). Both k_{cat} and K_m improved during the intermediate stages of closing but the majority of the effect on K_m occurred during the final stages of closing.

As the magnitude of the allosteric effect was on the same order as that of many natural allosteric enzymes, we next examined the biological effects of RG13. The switching activity was sufficient to result in an observable phenotype: maltose-dependent resistance to ampicillin (Table 1). *E. coli* cells expressing RG13 had a minimum inhibitory concentration (MIC) for ampicillin that was four-fold higher in the presence of 50 μM maltose. In contrast, the addition of the same concentration of sucrose or glucose to a plate did not affect the MIC. Thus, RG13 serves to couple the previously unrelated functions of ampicillin resistance and maltose concentration. *E. coli* cells expressing RG13 function as a growth/no growth sensor for maltose.

Significance

We have shown that two unrelated proteins can be systematically recombined in order to link their respective functions and create molecular switches. A combination of random circular permutation and random domain insertion enabled the creation of a MBP-BLA fusion geometry in which conformational changes induced upon maltose binding could propagate to the active site of BLA and increase BLA enzymatic activity up to 25-fold. The functional coupling of two proteins with no evolutionary or functional relationship is a powerful strategy for engineering novel molecular function. For example, coupling a ligand-binding protein and a protein with good signal transduction properties would result in the creation of a molecular sensor for the ligand. Furthermore, switches that establish connections between cellular components with no previous relationship can result in novel cellular circuitry and phenotypes. We envision, for example, that such switches might establish connections between molecular signatures of disease (e.g. abnormal concentrations of proteins, metabolites, signaling or other molecules) and functions that serve to treat the disease (e.g. delivery of drugs, modulation of signaling pathways or modulation of gene expression) and therefore possess selective therapeutic properties.

Supplemental Data

Supplemental data including experimental procedures and a figure are available at...

Acknowledgements

We thank David H. Gracias, Robert Schleif, and Denis Wirtz for helpful comments on the manuscript and Tom Mansell for constructing BLA(W208G). This work was supported by a grant from NIH.

References

1. Spencer, D.M., Wandless, T.J., Schreiber, S.L., and Crabtree, G.R. (1993). Controlling signal transduction with synthetic ligands. Science *262*, 1019-1024.
2. Picard, D. (2000). Posttranslational regulation of proteins by fusions to steroid-binding domains. Methods Enzymol *327*, 385-401.
3. Posey, K.L., and Gimble, F.S. (2002). Insertion of a reversible redox switch into a rare-cutting DNA endonuclease. Biochemistry *41*, 2184-2190.
4. Saghatelian, A., Guckian, K.M., Thayer, D.A., and Ghadiri, M.R. (2003). DNA detection and signal amplification via an engineered allosteric enzyme. J Am Chem Soc *125*, 344-345.
5. Brennan, C., Cristianson, K., Surowy, T., and Mandecki, W. (1994). Modulation of enzyme activity by antibody binding to an alkaline phosphatase-epitope hybrid protein. Protein Eng. *7*, 509-514.
6. Pelletier, J.N., Campbell-Valois, F.X., and Michnick, S.W. (1998). Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. Proc.Natl.Acad.Sci.USA *95*, 12141-12146.

7. Dueber, J.E., Yeh, B.J., Chak, K., and Lim, W.A. (2003). Reprogramming control of an allosteric signaling switch through modular recombination. Science *301*, 1904-1908.

8. Mootz, H.D., and Muir, T.W. (2002). Protein splicing triggered by a small molecule. J Am Chem Soc *124*, 9044-9045.

9. Guo, Z., Zhou, D., and Schultz, P.G. (2000). Designing small-molecule switches for protein-protein interactions. Science *288*, 2042-2045.

10. Gryczynski, U., and Schleif, R. (2004). A portable allosteric mechanism. Proteins *57*, 9-11.

11. Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (1999). Circular permutation and receptor insertion within green fluorescent proteins. Proc.Natl.Acad.Sci.U.S.A *96*, 11241-11246.

12. Tucker, C.L., and Fields, S. (2001). A yeast sensor of ligand binding. Nat Biotechnol *19*, 1042-1046.

13. Doi, N., and Yanagawa, H. (1999). Insertional gene fusion technology. FEBS Lett *457*, 1-4.

14. Radley, T.L., Markowska, A.I., Bettinger, B.T., Ha, J.H., and Loh, S.N. (2003). Allosteric switching by mutually exclusive folding of protein domains. J Mol Biol *332*, 529-536.

15. Guntas, G., and Ostermeier, M. (2004). Creation of an allosteric enzyme by domain insertion. J Mol Biol *336*, 263-273.

57

16. Buskirk, A.R., Ong, Y.C., Gartner, Z.J., and Liu, D.R. (2004). Directed evolution of ligand dependence: Small-molecule-activated protein splicing. Proc Natl Acad Sci U S A *101*, 10505-10510.

17. Ostermeier, M., and Benkovic, S.J. (2000). Evolution of protein function by domain swapping. Adv.Protein Chem. *55*, 29-77.

18. Apic, G., Gough, J., and Teichmann, S.A. (2001). Domain combinations in archaeal, eubacterial and eukaryotic proteomes. J Mol Biol *310*, 311-325.

19. Sharff, A.J., Rodseth, L.E., Spurlino, J.C., and Quiocho, F.A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotzxis. Biochemistry *31*, 10657-10663.

20. Graf, R., and Schachman, H.K. (1996). Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase. Proc.Natl.Acad.Sci.USA *93*, 11591-11596.

21. Ostermeier, M., and Benkovic, S.J. (2001). Construction of hybrid gene libraries involving the circular permutation of DNA. Biotechnol. Lett. *23*, 303-310.

22. Quiocho, F.A., Spurlino, J.C., and Rodseth, L.E. (1997). Extensive features of tight oligosaccharide binding revealed in high- resolution structures of the maltodextrin transport/chemosensory receptor. Structure *5*, 997-1015.

23. Maveyraud, L., Massova, I., Birck, C., Miyashita, K., Samama, J.-P., and Mobashery, S. (1996). Crystal structure of 6alpha-(hydroxymethyl)penicillanate complexed to the TEM-1 β-lactamase from Escherichia coli: evidence on the

mechanism of action of a novel inhibitor designed by a computer-aided process. J. Am. Chem. Soc. *118*, 7435-7440.

24. Evenas, J., Tugarinov, V., Skrynnikov, N.R., Goto, N.K., Muhandiram, R., and Kay, L.E. (2001). Ligand-induced structural changes to maltodextrin-binding protein as studied by solution NMR spectroscopy. J Mol Biol *309*, 961-974.

25. Sigal, I.S., DeGrado, W.F., Thomas, B.J., and Petteway, S.R., Jr. (1984). Purification and properties of thiol beta-lactamase. A mutant of pBR322 beta-lactamase in which the active site serine has been replaced with cysteine. J Biol Chem *259*, 5327-5332.

26. Christensen, H., Martin, M.T., and Waley, S.G. (1990). Beta-lactamases as fully efficient enzymes. Determination of all the rate constants in the acyl-enzyme mechanism. Biochem J *266*, 853-861.

27. Hall, J.A., Gehring, K., and Nikaido, H. (1997). Two modes of ligand binding in maltose-binding protein of *Eschericia coli*: correlation with the structure of ligands and the structure of binding protein. J Biol Chem *272*, 17605-17609.

28. Marvin, J.S., and Hellinga, H.W. (2001). Manipulation of ligand binding affinity by exploitation of conformational coupling. Nat Struct Biol *8*, 795-798.

29. Millet, O., Hudson, R.P., and Kay, L.E. (2003). The energetic cost of domain reorientation in maltose-binding protein as studied by NMR and fluorescence spectroscopy. Proc Natl Acad Sci U S A *100*, 12700-12705.

55

Table 1. Ampicillin resistance of *E. coli* cells in the presence and absence of maltose.

Expressed Protein	Minimum Inhibitory Concentration of Ampicillin (μg/ml)*	
	No maltose	50 μM maltose
none	4	4
RG13	128	512
BLA(W208G)†	32	32
BLA	≥2000	≥2000

*Conditions: DH5α-E cells on LB plates (with or without maltose) incubated at 37°C for 20 hours.

†A mutant of BLA with reduced activity.

Figure Legends

Figure 1. Creation of MBP-BLA molecular switches by non-homologous recombination. (A) The fragment of the BLA gene coding for the mature protein (codons 24-286) is flanked by sequences coding for a GSGGG linker (each of which contains a *BamHI* site). The fragment is excised by digestion with *BamHI* and cyclized by ligation under dilute DNA concentrations. A single, randomly-located double strand break is introduced by DNaseI digestion to create the circularly permuted library. This library is randomly inserted into plasmid pDIMC8-MBP containing the MBP gene (*malE*) under control of the *tac* promoter (*tacP/O*). The site for insertion in pDIMC8-MBP is created by introduction of a randomly located double-stranded break by digestion with dilute concentrations of DNaseI.
(B) Sequence of switch. Regions derived from MBP are shown in blue and those from BLA are shown in red. The GSGGG linker and a serine derived from the fusion of partial codons are shown in black. The number in parenthesis indicates the amino acid number of the starting proteins. The numbering system for MBP does not include the signal sequence. The numbering system for BLA does include the signal sequence and does not follow the consensus numbering system for β-lactamases.
(C) Structures of maltose-bound MBP [22] and BLA bound to an active-site inhibitor [23] oriented such that the fusion sites in RG13 are proximal. Dark blue, MBP[1-316]; yellow, BLA[227-286]; white, GSGGG linker; red, BLA[24-226]; green, MBP[319-370]; light blue, MBP deleted residues 317 and 318.

Figure 2. Switch activity in RG13 is specific to ligands of MBP and is reversible.

59

(A) The percent increase in the initial velocity of nitrocefin hydrolysis at 20 μM nitrocefin upon addition 5 mM of the indicated ligands (maltose, maltotriose and β-cyclodextrin) and non-ligands (sucrose, lactose and galactose).

(B) Reversible switching using competing ligand. During the enzymatic hydrolysis of nitrocefin, formation of product is monitored by absorbance at 486 nm. At time zero the reaction is started in 2 ml phosphate buffer (0.1 M) with 20 μM nitrocefin and 2.5 nM RG13. At the time indicated by the first arrow, 20 μl of 1 M maltose was added resulting in a 10-fold increase in the reaction rate. This maltose concentration is above the K_d for maltose but is subsaturating. At the time indicated by the second arrow, 230 μl of 10 mM β-cyclodextrin was added (final concentrations are 1.0 mM β-cyclodextrin and 8.9 μM maltose). Because RG13 has similar affinities for maltose and β-cyclodextrin but β-cyclodextrin is present at a >100-fold higher concentration, the β-cyclodextrin preferentially replaces the maltose bound to RG13 and the rate of reaction decreases to a level consistent with β-cyclodextrin's modest effect on nitrocefin hydrolysis.

(C) Reversible switching after dialysis. The initial rate of nitrocefin hydrolysis at 25 μM nitrocefin was measured at the indicated steps. Maltose was added to a final concentration of 5 mM.

Figure 3. Characterization of switches.

(A) Dissociation constants for maltose were determined in the absence (white bars) and presence (black bars) of saturating concentrations of carbenicillin. The apparent dissociation constants in the presence of subsaturating concentrations (25 μM) of nitrocefin (grey bars)

56

were also determined. The dissociation constants for maltose of MBP, MBP(I329W), MBP(I329W/A96W) (dashed line) reported by Marvin and Hellinga [28] are shown for comparison.

(B) – (D) Steady state kinetic parameters of nitrocefin hydrolysis for RG13, RG13(I329W) and RG13(I329W/A96W) in the presence (black bars) or absence (white bars) of saturating concentrations of maltose. Experimental conditions: 100 mM sodium phosphate buffer, pH 7.0, 25°C.

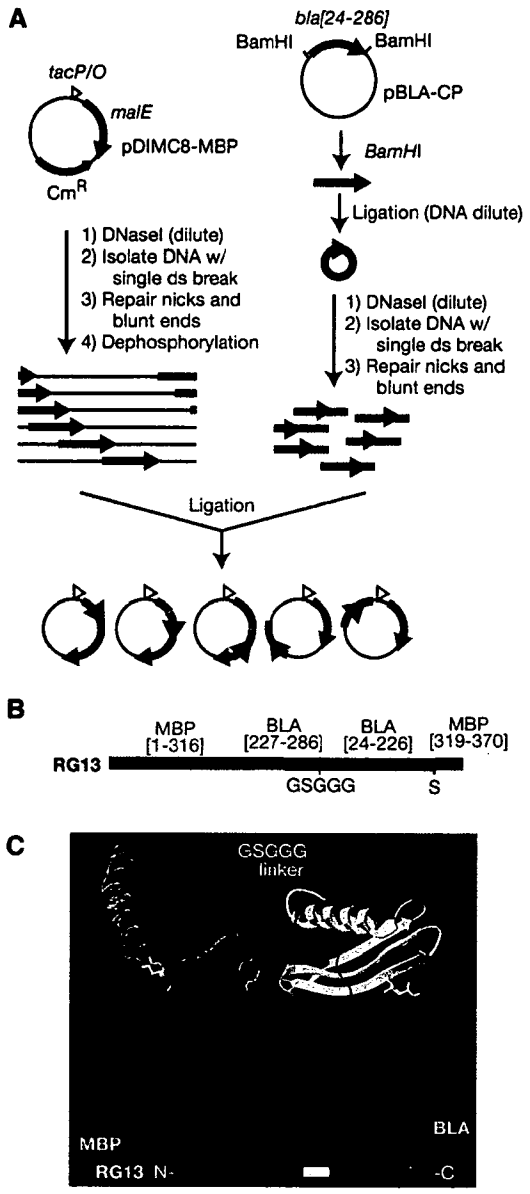
A
bla[24-286]
BamHI
BamHI
pBLA-CP
tacP/O
malE
pDIMC8-MBP
Cm^R
BamHI
Ligation (DNA dilute)
1) DNaseI (dilute)
2) Isolate DNA w/ single ds break
3) Repair nicks and blunt ends
4) Dephosphorylation
1) DNaseI (dilute)
2) Isolate DNA w/ single ds break
3) Repair nicks and blunt ends
Ligation
B
MBP [1-316]
BLA [227-286]
BLA [24-226]
MBP [319-370]
RG13
GSGGG
S
C
GSGGG linker
MBP
BLA
RG13 N-
-C

Figure 1

60

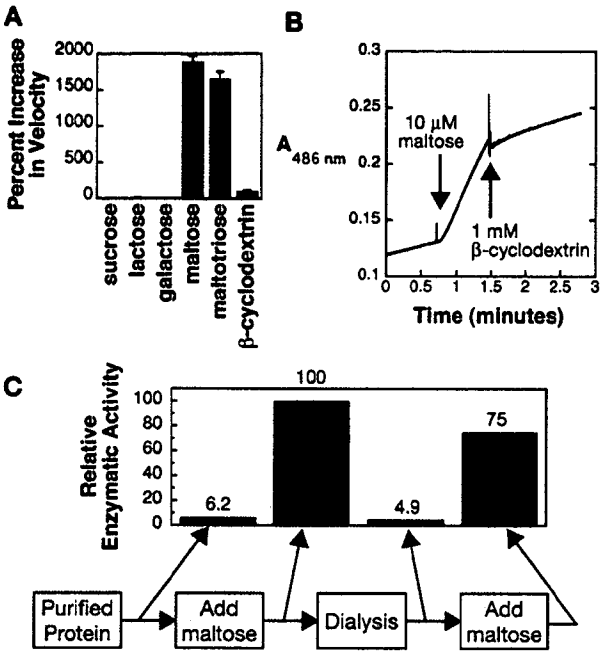
A
Percent Increase in Velocity
2000
1500
1000
500
0
sucrose
lactose
galactose
maltose
maltotriose
β-cyclodextrin
B
A 486 nm
0.3
0.25
0.2
0.15
0.1
10 μM maltose
1 mM β-cyclodextrin
0 0.5 1 1.5 2 2.5 3
Time (minutes)
C
Relative Enzymatic Activity
100
80
60
40
20
0
6.2
100
4.9
75
Purified Protein
Add maltose
Dialysis
Add maltose

Figure 2

61

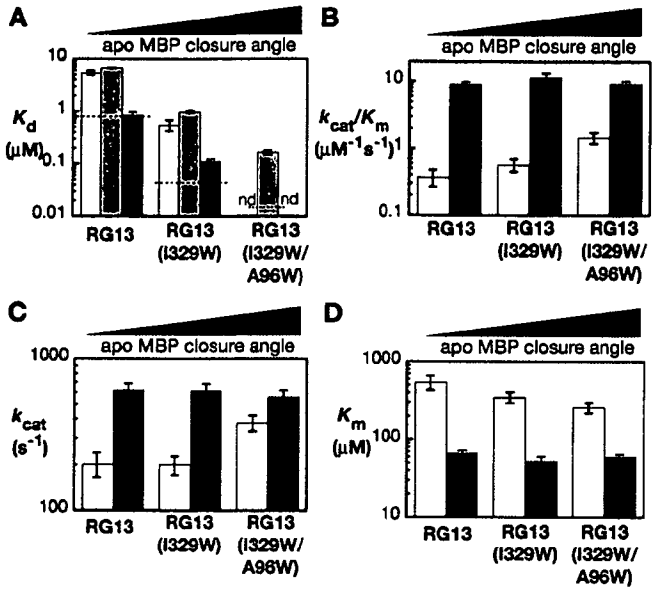
A
apo MBP closure angle
10
1
0.1
0.01
K_d (μM)
nd
nd
RG13
RG13 (I329W)
RG13 (I329W/ A96W)
B
apo MBP closure angle
10
1
0.1
k_{cat}/K_m ($\mu M^{-1}s^{-1}$)
RG13
RG13 (I329W)
RG13 (I329W/ A96W)
C
apo MBP closure angle
1000
100
k_{cat} (s^{-1})
RG13
RG13 (I329W)
RG13 (I329W/ A96W)
D
apo MBP closure angle
1000
100
10
K_m (μM)
RG13
RG13 (I329W)
RG13 (I329W/ A96W)

Figure 3

62

Supporting Data for Guntas, Mitchell, and Ostermeier.

Experimental Procedures

Supplemental Figure S1

Experimental Procedures

Materials

All restriction enzymes, T4 DNA ligase, T4 DNA polymerase, and calf intestinal phosphatase were purchased from New England Biolabs (Beverly, MA). pGEM T-vector cloning kit and Taq polymerase were purchased from Promega (Madison, WI). DNAseI was purchased from Roche Biochemicals (Indianapolis, IN). Qiaquick PCR purification kit and Qiaquick gel extraction kit were purchased from Qiagen (Valencia, CA). Popculture reagent, rLysozyme, benzonase nuclease, and His-tag protein purification kit were purchased from Novagen (Madison, WI). Oligonucleotides and Electromax DH5α-E electrocompetent cells were purchased from Invitrogen (Carlsbad, CA). Nitrocefin was purchased from Oxoid (Hampshire, UK). Maltotriose and β-cyclodextrin were purchased from Sigma (St. Louis, MO). Antibiotics, maltose, lactose, galactose and sucrose were purchased from Fisher Scientific (Pitsburgh, PA).

Random circular permutation.

The portion of the *bla* gene encoding the mature BLA was fused to a sequence coding for a GSGGG linker and containing a *BamHI* site by PCR amplification using the

forward primer 5'-TTTGCC**GGATCC**GGCGGTGGCCACCCAGAAACGCTGGTG-3' and the reverse primer 5'- GTCTGA**GGATCC**CCAATGCTTAATCAGTGA-3'. Portions of the primers encoding the GSGGG linker are underlined and the *BamHI* site is highlighted in bold. The PCR product was desalted using Qiaquick PCR purification kit and ligated to the pGEM T-vector to create plasmid pGEMT-BLA. One hundred and fifty μg of pGEMT-BLA was digested with 1000 units of *BamHI* and the DNA fragment that encodes BLA was gel purified using Qiaquick gel purification kit. Eighteen μg of this DNA was cyclized by ligation at 16°C for 18 hours in a reaction volume of 5.1 ml in the presence of ligase buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml BSA pH 7.5) and 600 Weiss units of T4 DNA ligase. After heat inactivation of the ligase, the concentrated reaction mixture was desalted and the circular DNA was purified by agarose gel electrophoresis using Qiaquick Gel Extraction kit. To introduce the random double stranded break, 8 μg of circular DNA was digested with 8 milliunits of DNAse I in the presence of 50 mM Tris-HCl, pH 7.4, 10 mM $MnCl_2$ and 50 μg/ml BSA in a total volume of 0.8 ml for 8 minutes. The reaction was quenched by the addition of EDTA to a concentration of 5 mM and the solution was desalted using a Qiaquick PCR purification column. Nicks and gaps were repaired by incubating at 12°C for 30 minutes in a total volume of 90 μl in the presence of T4 DNA polymerase (6 units) and T4 DNA ligase (12 Weiss units) in the presence of 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml BSA and 125 μM dNTPs. The DNA was desalted as before and the linear DNA (corresponding to the randomly circularly permuted *bla*) was

isolated from circular forms by agarose gel electrophoresis using the Qiaquick gel purification kit.

Random Domain Insertion.

Plasmid pDIM-C8MalE has the *malE* gene encoding MBP under the IPTG inducible *tac* promoter. Introduction of a random double stranded breaks (one per molecule of pDIM-C8MalE) was performed as described [1]. One hundred ng of randomly linearized plasmid pDIMC8-MalE was ligated to 85 ng of randomly circularly permuted BLA fragment (5:1 insert/vector molar ratio) in a reaction volume of 15 μl. The ligation was carried out at 22°C overnight in the presence of ligase buffer and 45 Weiss units T4 DNA ligase. After ethanol precipitation, the ligated DNA was transformed into Electromax DH5α-E electrocompetent cells by performing ten electroporations of 40 μl cells each. Cells were plated on two 245x245 mm LB agar plates supplemented with 50 μg/ml chloramphenicol and incubated at 37°C overnight. The naïve library was recovered from the large plates and stored in frozen aliquots as described [2].

Library Selection and Screening

The naïve library was plated on LB agar plates supplemented with 200 μg/ml ampicillin and 50 mM maltose and incubated at 37°C overnight. From these plates, 1056 colonies were picked to inoculate 1 ml LB media (supplemented with 50 μg/ml chloramphenicol and 1 mM IPTG) in 96-well format. After incubation overnight at 37°C, each culture was

lysed using 0.1 ml Popculture reagent, 40 units of rLysozyme, and 2.5 units of benzonase nuclease. Lysates were centrifuged to pellet the insoluble material and the soluble fractions were assayed in 96-well format for nitrocefin hydrolysis in the presence or absence of 5 mM maltose using a colorimetric assay for nitrocefin hydrolysis [3]. The assays were carried out at room temperature using the Spectramax-384 Plus microplate reader (Molecular Devices) in the presence of 100 mM sodium phosphate buffer and 50 μM nitrocefin in a 200 μl reaction volume. Clones whose lysates exhibited a greater than 2-fold increase in the rate of nitrocefin hydrolysis were recultured and their lysates assayed again to verify the effect.

Protein Modifications and Mutagenesis

A $GGSGH_9$ sequence was appended to the sequence of RG13 by PCR amplification with the appropriate primers. The PCR product was cloned between *Nde*I and *Xho*I sites of pET24b (Novagen) to create pET24b-RG13. Mutations I329W and A96W were introduced into pET24b-RG13 by a combination of overlap extension PCR and Quickchange mutageneis.

Protein Purification

One liter LB media containing 50 μg/ml kanamycin was inoculated with 2% overnight culture and shaken at 37 °C. The culture was induced with 1 mM IPTG when the OD_{600} reached 0.5 and incubated at 22 °C for 16 hours. Pelleted cells were resuspended in 20 ml

binding buffer supplied by the His-tag protein purification kit (Novagen, Madison, WI) and lysed by French press. The soluble fraction was recovered and the protein was purified using the protein purification kit. Eluted protein was dialyzed at 4 °C against three liters of 100 mM sodium chloride, 50 mM sodium phosphate buffer overnight followed by dialysis against one liter of the same buffer with 20% glycerol for four hours. Protein was stored in aliquots at –80°C. RG13 and RG13(I329W) were purified as described above. To improve the yield of RG13(A96W/I329W), 10 mM maltose was added to the culture at induction. RG13(A96W/I329W) was dialyzed more extensively after purification and complete removal of maltose was verified by enzymatic assay on successive rounds of dialysis in the presence and absence of maltose. The purities of the proteins were estimated by coomassie blue staining of SDS-PAGE gels. The purities of RG13, RG13(I329W), and RG13(A96W/I329W) were greater than 98%, 95% and 97%, respectively. The extinction coefficients of RG13, RG13(I329W), and RG13(A96W/I329W) at 280 nm were calculated [4] to be 126,000; 120,500 and 116,100 Abs M^{-1} cm^{-1}, respectively.

Steady State Kinetics

All kinetic assays were performed at 25°C in the presence of 100 mM sodium phosphate buffer, pH 7.0. Ten µl of enzyme stock was added to 1.59 ml buffer (containing the saccharide, if desired). After incubation for 30 seconds, 0.4 ml of 5x substrate was added and the absorbance at the appropriate wavelength was recorded using the Cary50 UV-VIS spectrophotometer. The wavelength monitored was 486 nm, 240 nm, and 232 nm for

nitrocefin, carbenicillin, and ampicillin respectively. From the initial rate of reaction the kinetic constants were determined using Eadie-Hofstee plots. In the absence of maltose, the time course of the reaction for RG13, RG13(I329W), and RG13(A96W/I329W) displayed a slight lag in the reaction rate that became more pronounced at higher substrate concentrations. The rate data was consistent with a small hysteretic effect [5] and not substrate inhibition as preincubation of the enzyme with the substrate for one minute prevented the lag from occurring upon addition of more substrate. Therefore, the steady state parameters for nitrocefin hydrolysis in the absence of maltose were determined by measuring the rate at 1-2 minutes (well after the lag) and correcting the substrate concentration by subtracting the amount of substrate hydrolyzed. In all cases the extent of reaction at the point the rate was measured was less than 25%. In the presence of maltose, no lag was observed.

Maltose Affinity

Maltose affinity for RG13 (in presence and absence of 10 mM carbenicillin) and RG13(I329W) (in the absence of substrate) was determined using intrinsic protein fluorescence measured on a Photon Technology QuantaMaster QM-4 spectrofluorometer. Fluorescence spectra were obtained at 25 °C at different concentrations of maltose in 50 mM sodium phosphate buffer, pH 7.0, containing 100 mM sodium chloride. The protein concentration was 50-100 nM. Excitation was at 280 nm. The quenching in fluorescence intensity at 341 nm caused by maltose was used in Eadie-Hofstee equivalent plots to

determine K_d using the following equation: $\Delta F = \Delta F_{max} - K_d \frac{\Delta F}{[L]}$; where ΔF is the change in fluorescence intensity, ΔF_{max} is the difference in fluorescence between no maltose and saturating amounts of maltose and [L] is the maltose concentration. The fluorescence quenching of RG13(I329W/A96W) upon addition of maltose was insufficient to accurately determine a K_d by this method. The dissociation constant for maltose and RG13(I329W) in the presence of saturating carbenicillin (2 mM) was determined by measuring the initial rate of carbenicillin hydrolysis as a function of maltose concentration. The apparent dissociation constant in the presence of subsaturating concentrations of nitrocefin (25 μM) for all three proteins was determined by measuring the initial rate of nitrocefin hydrolysis as a function of maltose concentration.

***In Vivo* Characterization of Switches.**

Overnight inoculums of DH5α-E cells expressing RG13, BLA or BLA(W208G) were diluted into LB media and plated on LB plates, either in the absence or presence of 50 μM maltose, in the presence of increasing amounts of ampicillin. Ampicillin was present in the plates at the following concentrations: 0, 2, 4, 8, 16, 32, 64, 128, 256, 512, and 2000 μg/ml. Cells were plated at approximately 1000 CFU (no antibiotic) per plate. The plates were incubated at 37 °C for 20 hours. The minimum inhibitory concentration (MIC) was defined as the lowest ampicillin concentration at which no colonies were present, or that at which the number of colonies present was <1% of the number of colonies at the next lowest level of ampicillin.

References

1. Guntas, G., and Ostermeier, M. (2004). Creation of an allosteric enzyme by domain insertion. J Mol Biol *336*, 263-273.

2. Ostermeier, M., Nixon, A.E., Shim, J.H., and Benkovic, S.J. (1999). Combinatorial protein engineering by incremental truncation. Proc.Natl.Acad.Sci.USA *96*, 3562-3567.

3. O'Callaghan, C.H., Morris, A., Kirby, S.M., and Shingler, A.H. (1972). Novel method for detection of beta-lactamases by using a chromogenic cephalosporin substrate. Antimicrob Agents Chemother *1*, 283-288.

4. Gill, S.C., and von Hippel, P.H. (1989). Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem *182*, 319-326.

5. Neet, K.E., and Ainslie, G.R., Jr. (1980). Hysteretic enzymes. Methods Enzymol *64*, 192-226.

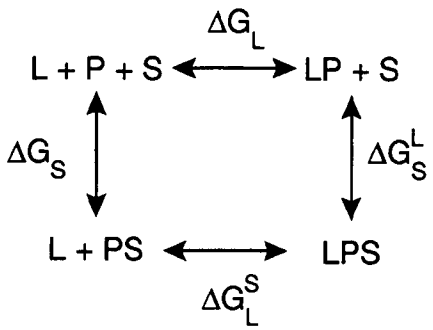

Figure S1. Coupling of ligand and substrate binding. The change in free energy upon protein (P) binding ligand (L) and substrate (S) is the same whether ligand or substrate binds first. Adding the free energy changes of the two different paths from L+P+S to LPS we see that $\Delta G_L + \Delta G_S^L = \Delta G_S + \Delta G_L^S$ since the total free energy change is path independent. By rearranging this equation to $\Delta G_L - \Delta G_L^S = \Delta G_S - \Delta G_S^L$ we see that the left hand side represents the effect that the presence of bound substrate has on ligand binding and the right hand side represents the effect that the presence of bound ligand has on substrate binding. The effects must be equal.